An Introduction to Macroeconomics

An Introduction to Macroeconomics, Second Edition

A HETERODOX APPROACH TO ECONOMIC ANALYSIS

Edited by

Louis-Philippe Rochon
Full Professor, Laurentian University, Canada; Editor-in-Chief, Review of Political Economy *and Founding Editor Emeritus,* Review of Keynesian Economics

Sergio Rossi
Full Professor of Economics, University of Fribourg, Switzerland

Edward Elgar PUBLISHING

Cheltenham, UK • Northampton, MA, USA

© Louis-Philippe Rochon and Sergio Rossi 2021

Cover image: From the private collection of Louis-Philippe Rochon.
"*Maybe Wednesday no.6*" (2018) by Florence Victor.
Oil and graphite on canvas. 48x48

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2020952391

ISBN 978 1 78990 114 6 (cased)
ISBN 978 1 78990 116 0 (paperback)
ISBN 978 1 78990 115 3 (eBook)

Typeset by Servis Filmsetting Ltd, Stockport, Cheshire

BRITISH LIBRARY

Contents in brief

List of contributors xiii
Acknowledgements xxi

Introduction: the urgent need for a heterodox approach to economic analysis 1
Louis-Philippe Rochon and Sergio Rossi

PART I ECONOMICS, ECONOMIC ANALYSIS AND ECONOMIC SYSTEMS

1 What is economics? 25
Louis-Philippe Rochon and Sergio Rossi

2 The state of macroeconomics 51
Louis-Philippe Rochon and Sergio Rossi

3 The history of economic theories 80
Heinrich Bortis

4 Monetary economies of production 122
Louis-Philippe Rochon

PART II MONEY, BANKS AND FINANCIAL ACTIVITIES

5 Money and banking 151
Marc Lavoie and Mario Seccareccia

6 The financial system 176
Jan Toporowski

7 Central banking and monetary policy 200
Louis-Philippe Rochon and Sergio Rossi

PART III THE MACROECONOMICS OF THE SHORT AND LONG RUN

8 Theories of consumption 233
Stavros A. Drakopoulos

9 Theories of investment 268
Thomas Dallery

10 Aggregate demand 307
Jesper Jespersen

11 Inflation and unemployment 331
Alvaro Cencini and Sergio Rossi

12 The role of fiscal policy 355
Malcolm Sawyer

13 Economic growth and development 376
Mark Setterfield

14 Wealth distribution 402
Omar Hamouda

PART IV INTERNATIONAL ECONOMY

15 International trade and development 437
Robert A. Blecker

16 Balance-of-payments constrained growth 469
John McCombie and Nat Tharnpanich

17 European monetary union 494
Sergio Rossi

PART V RECENT TRENDS

18 Financialization 517
Gerald A. Epstein

19 Imbalances and crises 540
Robert Guttmann

20 Sustainable development 566
Richard P.F. Holt

21 Conclusion: do we need microfoundations for macroeconomics? 593
John King

Answers to the exam questions 618

Index 633

Full contents

List of contributors xiii
Acknowledgements xxi

Introduction: the urgent need for a heterodox approach to economic analysis **1**
Louis-Philippe Rochon and Sergio Rossi

PART I ECONOMICS, ECONOMIC ANALYSIS AND ECONOMIC SYSTEMS

1 What is economics? **25**
Louis-Philippe Rochon and Sergio Rossi
Introduction 25
The role of ideology in economics 31
Is economics a science? 34
The use of models and of mathematics 37
Economics and the social sciences 40
What then is economics? 43
Micro- versus macroeconomics 45
A portrait of Adam Smith (1723–90) 48

2 The state of macroeconomics **51**
Louis-Philippe Rochon and Sergio Rossi
Why are these topics important? 51
The traditional mainstream view 58
Rethinking DSGE models? 62
An alternative perspective: rejecting the models? 66
Conclusion 69
A portrait of Piero Sraffa (1898–1983) 76

3 The history of economic theories **80**
Heinrich Bortis
Why are these topics important? 82
Two broad groups of economic theories 82
The history of economics 84

The history of political economy 92
Neoclassical–Walrasian economics and classical–Keynesian political economy assessed 113
A portrait of David Ricardo (1772–1823) 118

4 Monetary economies of production 122
Louis-Philippe Rochon
Why are these topics important? 123
The neoclassical/mainstream view 123
The heterodox view 128
Conclusion 143
A portrait of John Maynard Keynes (1883–1946) 145

PART II MONEY, BANKS AND FINANCIAL ACTIVITIES

5 Money and banking 151
Marc Lavoie and Mario Seccareccia
Why are these topics important? 152
The traditional mainstream view 152
The heterodox perspective 156
Understanding the heterodox approach to banks and the modern payments system from a simple balance sheet perspective 161
Concluding remarks 169
A portrait of Alain Parguez (1940–) 172

6 The financial system 176
Jan Toporowski
Why are these topics important? 177
The traditional mainstream view 181
The heterodox perspective 185
The economic consequences of long-term finance 190
Finance in Keynes's analysis 193
Modern finance 194
Conclusion 195
A portrait of Hyman Philip Minsky (1919–96) 197

7 Central banking and monetary policy 200
Louis-Philippe Rochon and Sergio Rossi
Why are these topics important? 201
The mainstream perspective 202
The heterodox perspective 213

Conclusion 224
A portrait of Alfred S. Eichner (1937–88) 227

PART III THE MACROECONOMICS OF THE SHORT AND LONG RUN

8 Theories of consumption 233
Stavros A. Drakopoulos
Why are these topics important? 234
The mainstream view: intertemporal choice and consumption function 235
Keynes's approach to consumption 238
Empirical testing of consumption functions 244
Mainstream consumption theories 247
The heterodox perspective 253
Policy implications and concluding remarks 260
A portrait of James Stemble Duesenberry (1918–2009) 263

9 Theories of investment 268
Thomas Dallery
Why are these topics important? 269
The traditional mainstream view 271
Keynes's approach to investment 275
The heterodox perspective 288
Empirical testing of investment functions 297
Concluding remarks 300
A portrait of Henry Roy Forbes Harrod (1900–78) 303

10 Aggregate demand 307
Jesper Jespersen
Why are these topics important? 308
From statistical concepts to macroeconomic theory 310
The neoclassical theory of aggregate demand 312
The Keynesian theory of aggregate demand and output 315
The income multiplier 321
Expected aggregate demand and supply: effective demand 323
Demand management policies 325
A portrait of Richard Kahn (1905–89) 328

11 Inflation and unemployment 331
Alvaro Cencini and Sergio Rossi
Why are these topics important? 332

Inflation 333
Unemployment 340
Towards a monetary macroeconomic analysis of inflation and unemployment 343
A portrait of Bernard Schmitt (1929–2014) 352

12 The role of fiscal policy 355
Malcolm Sawyer
Why are these topics important? 356
Is there a need for fiscal policy? 356
The traditional mainstream view 358
The post-Keynesian perspective 360
The role of automatic stabilizers 363
Functional finance and the post-Keynesian approach to fiscal policy 365
Concluding remarks 371
Appendix 372
A portrait of Michał Kalecki (1899–1970) 373

13 Economic growth and development 376
Mark Setterfield
Why are these topics important? 377
The traditional mainstream view: supply-led growth 382
The Keynesian view: demand-led growth 384
Sources of demand in the long run 386
Keynesian growth theory 387
Properties of Keynesian growth theory 392
Conclusion 396
A portrait of Nicholas Kaldor (1908–86) 398

14 Wealth distribution 402
Omar Hamouda
Why are these topics important? 403
Some introductory remarks 404
Mainstream economic theory of wealth distribution 405
States of income distribution: a description 413
Heterodox perspectives 419
Some concluding remarks 427
A portrait of Karl Marx (1818–83) 430

PART IV INTERNATIONAL ECONOMY

15 International trade and development **437**
Robert A. Blecker
Why are these topics important? 438
The orthodox approach: the theory of comparative advantage 440
The heterodox alternative: imbalanced trade, unemployment and absolute competitive advantages 447
Long-run development and infant-industry protection 451
Trade liberalization, trade agreements and trade wars 453
Manufactured exports, the fallacy of composition and global value chains 460
Conclusion 462
A portrait of Joan Robinson (1903–83) 466

16 Balance-of-payments constrained growth **469**
John McCombie and Nat Tharnpanich
Why are these topics important? 470
The traditional mainstream view 470
The heterodox alternative 472
Concluding comments 488
A portrait of Anthony Philip Thirlwall (1941–) 490

17 European monetary union **494**
Sergio Rossi
Why are these topics important? 495
The mainstream perspective 495
The heterodox perspective 502
A portrait of Robert Triffin (1911–93) 512

PART V RECENT TRENDS

18 Financialization **517**
Gerald A. Epstein
Why are these topics important? 518
What is financialization? 519
How old is financialization? 521
Dimensions of financialization 521
Impacts of financialization 528
Conclusion 533
A portrait of Karl Paul Polanyi (1886–1964) 536

19 Imbalances and crises **540**
Robert Guttmann
Why are these topics important? 541
The mainstream view of a self-adjusting economy 541
The heterodox alternative 542
Growth and distribution 546
Cyclical growth dynamics 547
External imbalances and adjustments 553
Concluding remarks 560
A portrait of Mikhail Tugan-Baranovsky (1865–1919) 563

20 Sustainable development **566**
Richard P.F. Holt
Why are these topics important? 567
The neoclassical model of economic growth and development 568
The debate over 'strong' and 'weak' sustainability 569
Growth versus development 573
Heterodox economics and true economic development 575
Investments for sustainable development 579
Measuring a new standard of living for sustainable development 582
Conclusion 585
A portrait of Amartya Sen (1933–) 589

21 Conclusion: do we need microfoundations for macroeconomics? **593**
John King
Why are these topics important? 594
The mainstream perspective 595
A heterodox critique 602
Why it all matters 612
A portrait of Robert Skidelsky (1939–) 615

Answers to the exam questions 618

Index 633

Contributors

The editors

Louis-Philippe Rochon is Full Professor of Economics at Laurentian University, Canada, where he has been teaching since 1994. Before that, he taught at Kalamazoo College, in Michigan, United States. He obtained his doctorate from the New School for Social Research, in 1998, earning him the Frieda Wunderlich Award for Outstanding Dissertation, for his dissertation on endogenous money and post-Keynesian economics. Since January 2019, he has been the editor of the *Review of Political Economy*, and is now its Editor-in-Chief. He is also the founder and past editor (now emeritus) of the *Review of Keynesian Economics*. He has been guest editor for the *Journal of Post Keynesian Economics*, the *International Journal of Pluralism and Economics Education*, the *European Journal of Economic and Social Systems*, the *International Journal of Political Economy*, and the *Journal of Banking, Finance and Sustainable Development*. He has published on monetary theory and policy, post-Keynesian economics and fiscal policy. He is on the editorial board of *Ola Financiera*, *International Journal of Political Economy*, the *European Journal of Economics and Economic Policies: Intervention*, *Problemas del Desarrollo*, *Cuestiones Económicas* (Central Bank of Ecuador), and *Credit and Money* (Central Bank of Poland). He is the editor of the following book series: the *Elgar Series in Central Banking and Monetary Policy*, *Heterodox Undergraduate Introductions Series,* and *New Directions in Post-Keynesian Economics*. His recent books include *Employment in the Age of Austerity* (Edward Elgar Publishing, 2020), as well as two volumes honouring the work of Marc Lavoie and Mario Seccareccia (*Credit, Money and Crises in Post-Keynesian Economics*, Edward Elgar, 2020; and *Economic Growth and Macroeconomic Stabilization Policies in Post-Keynesian Economics*, Edward Elgar, 2020). He has been a visiting professor or visiting scholar in Australia, Brazil, France, Italy, Mexico, Poland, South Africa, and the United States, and has further lectured in China, Colombia, Ecuador, Italy, Japan, Kyrgyzstan, and Peru. He is the author of some 150 articles in peer-reviewed journals and books, and has written or edited close to 30 books. He has received grants from the Social Sciences and Humanities Research Council in Canada (SSHRC), the Ford Foundation and the Mott Foundation, among other places.

Sergio Rossi is Full Professor of Economics at the University of Fribourg, Switzerland, where he has held the Chair of Macroeconomics and Monetary Economics since 2005. He is a member of the Council of the Jean-Monnet Foundation for Europe at the University of Lausanne, Switzerland, and a blogger for the newspaper *Le Temps*. Since 2009 he has been a columnist for *Plusvalore*, an economics programme broadcast on Swiss Radio. After obtaining his PhD degree in Political Economy at the University of Fribourg (1996), where he was awarded the Vigener Prize for the best PhD dissertation in that year, he carried on his research work on monetary macroeconomics, and was honoured with two awards by the Committee of Vice-Chancellors and Principals of the United Kingdom for his PhD degree in Economics at University College London (2000). He has been Visiting Professor at the Centre for Banking Studies in Lugano (2000–2007) and at various universities in Europe. His research interests focus on macroeconomic analysis, particularly as regards national and international monetary and financial issues. He has authored and edited around 30 books (among them an encyclopedia of central banking and an encyclopedia of post-Keynesian economics), and is regularly invited to television and radio talk-shows discussing contemporary economic issues at national and international level. He has also published many peer-reviewed articles on monetary macroeconomics in *L'Actualité Economique: Revue d'analyse économique, Cambridge Journal of Regions, Economy and Society, China–USA Business Review, Cuadernos de Economía, European Journal of Economics and Economic Policies: Intervention, European Journal of Economic and Social Systems, Iberian Journal of the History of Economic Thought, International Journal of Monetary Economics and Finance, International Journal of Pluralism and Economics Education, International Journal of Political Economy, International Journal of Trade and Global Markets, International Review of Applied Economics, Investigación Económica, Journal of Asian Economics, Journal of Philosophical Economics, Journal of Post Keynesian Economics, Ola Financiera, Papers in Political Economy, Public Choice, Review of Keynesian Economics, Review of Political Economy, Revista de Economía Crítica, Studi Economici, Studi e Note di Economia* and in the *Swiss Journal of Economics and Statistics*. He is a member of the editorial board of *Cogent Economics and Finance, International Journal of Monetary Economics and Finance, Review of Keynesian Economics* and *Review of Political Economy*. In 2012, *L'Hebdo* included him in the 100 most prominent persons in French-speaking Switzerland. He also features in the 2015, 2017 and 2019 rankings of the most influential economists in Switzerland established by the *Neue Zürcher Zeitung*.

The contributors

Gustavo Bhering is a postdoctoral researcher in the Economics Department at Universidade Federal do Rio de Janeiro (UFRJ), where he is part of the political economy research group. He holds a bachelor's degree in Economic Sciences from the Pontifical Catholic University of Rio de Janeiro (2010), a master's degree in Economics from UFRJ (2013) and a PhD in Economics from UFRJ (2017). His PhD thesis on a reinterpretation of Ricardo's theory of foreign trade based on Sraffa (1930) was awarded the Pierangelo Garegnani Prize 2018. Bhering's research areas are in the Classical political-economy approach to value and distribution (based on Sraffa's interpretation), macroeconomics, growth, balance-of-payments constraint, and foreign trade.

Robert A. Blecker is Professor of Economics at American University (AU), Washington, DC, United States, Affiliated Faculty of AU's School of International Service and Center for Latin American and Latino Studies, and a Fellow of the Forum for Macroeconomics and Macroeconomic Policy. His most recent book is *Heterodox Macroeconomics: Models of Demand, Distribution and Growth* (co-authored with Mark Setterfield, Edward Elgar Publishing, 2019). His previous books include *Fundamentals of US Foreign Trade Policy: Economics, Politics, Laws, and Issues*, 2nd edition (co-authored with Stephen D. Cohen and Peter D. Whitney, Westview, 2003) and *Taming Global Finance: A Better Architecture for Growth and Equity* (Economic Policy Institute, 1999). His research includes work on post-Keynesian models of open economies, international trade theory and policy, economic integration in North America, global imbalances and the United States trade deficit, the Mexican economy, North–South trade and export-led growth.

Heinrich Bortis has been Professor Emeritus at the University of Fribourg, Switzerland, since 2015, where he was Full Professor of Political Economy from 1980. He has a PhD in Economics from the University of Cambridge, United Kingdom (Churchill College, 1977). His supervisors were Phyllis Deane and Nicholas Kaldor (thesis supervisor). Since 1998 he has been a Life Member of Clare Hall, University of Cambridge, and since 2013, a member of the German Keynes Society. His work is on fundamental issues in economic theory, specifically on the synthesis of Keynesian and Classical political economy. His publications include: *Institutions, Behaviour and Economic Theory – A Contribution to Classical-Keynesian Political Economy* (Cambridge University Press, 1997), 'Keynes and the Classics – notes on the monetary theory of production', in *Modern Theories of Money* (Edward Elgar Publishing, 2003) and 'Post-Keynesian principles and policies' in *Oxford Handbook of Post-Keynesian Economics* (Oxford University Press, 2013).

Alvaro Cencini is Emeritus Professor of Economics at the University of Lugano, Switzerland, where he held the Chair of Monetary Economics. He has a DPhil. degree in Economics from the University of Fribourg, Switzerland and a PhD degree in Economics of the London School of Economics, United Kingdom. His research work is in the field of national and international macroeconomics and he is presently working on a book about Bernard Schmitt's quantum macroeconomic analysis to be published by Routledge in 2021. Among his main publications are: *Time and the Macroeconomic Analysis of Income* (Pinter Publishers and St Martin's Press, 1984; Bloomsbury, 2013), *Money, Income and Time* (Pinter Publishers, 1988; Bloomsbury, 2013), *Monetary Macroeconomics: A New Approach* (Routledge, 2001, reprinted 2014), *Macroeconomic Foundations of Macroeconomics* (Routledge, 2005, reprinted 2007), *Economic and Financial Crises* (Palgrave Macmillan, 2015, co-authored with Sergio Rossi). He is also the author of several contributions to books and peer-reviewed journals. See also www.quantum-macroeconomics.info.

Thomas Dallery obtained a PhD at the University of Lille (France), and is currently Assistant Professor at the Université du Littoral Côte d'Opale (France). His main research areas deal with financialization (from both a microeconomic and a macroeconomic perspective) and fiscal policy (Keynesian multipliers and public debt sustainability). Notably, he contributed to research work on the cost of capital in French non-financial firms where shareholders' pressures are proven to lead to a decline in investment. His work has been published in the *Review of Radical Political Economy*, the *Review of Political Economy*, *Review of Keynesian Economics*, *Journal of Post Keynesian Economics*, and *Metroeconomica*, among others.

Stavros A. Drakopoulos obtained a BA in Economics from the Economics University of Athens, Greece, and an MSc and PhD in economics from the University of Stirling, United Kingdom. He was a Lecturer at the University of Glasgow (1988–1989), and at the University of Aberdeen (1990–1996), then Assistant Professor (1996–2000) and Associate Professor (2000–2005) at the National and Kapodistrian University of Athens. He is currently a Full Professor of Economics in the Department of Philosophy and History of Science of the National and Kapodistrian University of Athens (since September 2005). His research interests include history and methodology of economics, labour economics, and the economics of subjective well-being. He has published 50 peer-reviewed papers, five books, 13 chapters in edited books and encyclopaedias including books published by Routledge, Edward Elgar Publishing and Springer.

Gerald A.Epstein is Professor of Economics and a founding Co-Director of the Political Economy Research Institute at the University of Massachusetts

Amherst, United States. He has published articles on many topics including financialization, financial crises and regulation, the political economy of capital flows, and the political economy of central banking and financial institutions. He has worked with numerous United Nations agencies including the International Labour Organization and the United Nations Conference on Trade and Development. His most recent books are: *The Political Economy of International Finance in an Age of Inequality: Soft Currencies, Hard Landings* (Edward Elgar Publishing, 2018), *What's Wrong with Modern Money Theory: A Policy Critique* (Palgrave, 2019) and *The Political Economy of Central Banking: Contested Control and the Power of Finance* (Edward Elgar Publishing, 2019).

Robert Guttmann is Augustus B. Weller Professor of Economics in the Department of Economics at Hofstra University (New York, United States) and Professor Emeritus at the University Sorbonne Paris Nord (France). He studied in Vienna (Austria) and the University of Wisconsin-Madison (United States) before obtaining his PhD in Economics at the University of Greenwich in London, United Kingdom, in 1979. He won Distinguished Teacher of the Year awards at Hofstra in 1989, 2004 and 2012. He teaches on international economics, monetary economics, economic integration in the European Union, and public finance. Widely published in monetary theory as well as money and banking, his latest books are *Finance-Led Capitalism: Shadow Banking, Re-Regulation and the Future of Global Markets* (Palgrave Macmillan, 2016) and *Eco-Capitalism: Carbon Money, Climate Finance and Sustainable Development* (Palgrave Macmillan, 2018). He is currently working on another book, provisionally entitled *Multi-Polar Capitalism: The End of the Dollar Standard* and scheduled for publication in early 2021.

Omar Hamouda is Professor of Economics at Glendon College, York University, Canada, and has research expertise in macroeconomics and monetary economics, with a particular interest in and contributions to the theories of Hicks and Keynes. He is also well versed in the history of economic thought and in economic development. His book *Money, Investment and Consumption: Keynes's Macroeconomics Rethought* (Edward Elgar Publishing, 2009) is a rethinking of macroeconomics in light of Keynes's ideas on financial crises. He is the author or co-author of numerous books, including: *John R. Hicks: The Economist's Economist* (Basil Blackwell, 1993), *Probability in Economics* (with J.C.R. Rowley, Routledge, 1996) and *Verification in Economics and History: A Sequel to 'Scientification'* (with Betsy Price, Routledge, 2011). He has published many articles in peer-reviewed journals. He is editor of the *Journal of Income Distribution*.

Richard P.F. Holt is an award-winning writer. He has authored, co-authored or edited a number of books, two of which were named Choice

Outstanding Academic Book Title for the years 2005 and 2010. He recently edited a volume of selected letters by John Kenneth Galbraith for Cambridge University Press. He has also published over 60 articles and book reviews in a variety of academic journals, along with serving on different editorial boards. His research areas include environmental and ecological economics, post-Keynesian economics, history of economic thought and complexity economics. Presently, he is writing a book on the post-war conservative and liberal battles in the United States, for Yale University Press. He is a Professor of Economics at Southern Oregon University, United States.

Jesper Jespersen is a Professor Emeritus of Economics at Roskilde University, Denmark. He holds a PhD in Economics from the European University Institute, Florence, Italy and a Dr Scient. Adm. in Macroeconomic Methodology, Roskilde University. He is by-fellow at Churchill College, University of Cambridge, United Kingdom and has been Visiting Professor at the University of Bourgogne, Dijon, France. He is a member of the Post-Keynesian Society. He has published several books and peer-reviewed papers in Danish, English and Italian. Among his latest books are: *Handbook of Macroeconomic Methodology* (edited with Victoria Chick and Bert Tieben, Routledge, 2020), *The General Theory and Keynes for the 21st Century* (edited with Sheila Dow and Geoff Tily, Edward Elgar Publishing, 2018), and *Money, Method and Contemporary Post-Keynesian Economics* (edited with Sheila Dow and Geoff Tily, Edward Elgar Publishing, 2018).

John King is Emeritus Professor at La Trobe University and Honorary Professor at Federation University Australia. His principal research interests are in the history of heterodox economic thought, especially Marxian political economy and post-Keynesian economics. His recent publications include *The Distribution of Wealth* (with Michael Schneider and Mike Pottenger, Edward Elgar Publishing, 2016), *A History of American Economic Thought* (with Samuel Barbour and James Cicarelli, Routledge, 2018) and *The Alternative Austrian Economics* (Edward Elgar Publishing, 2019), dealing with socialist economic thought in Austria between 1900 and the present day.

Marc Lavoie is Professor Emeritus at the University Sorbonne Paris Nord (formerly Paris 13), France and at the University of Ottawa, Canada, where he taught for 38 years. He is a Research Fellow at the Macroeconomic Research Institute and at the Forum for Macroeconomics and Macroeconomic Policies of the Hans Böckler Foundation in Düsseldorf, and a Research Associate at the Broadbent Institute in Toronto. Lavoie has published ten books and nearly 250 refereed articles or book chapters, mostly in macroeconomics. He is best known for his book with Wynne Godley, *Monetary Economics* (Palgrave Macmillan, 2007), which is considered a must-read for

users of the stock-flow consistent approach. His latest book, *Post-Keynesian Economics: New Foundations* (Edward Elgar Publishing, 2014), received the 2017 Myrdal Prize from the European Association of Evolutionary Political Economy. He is a co-editor of two academic journals and is on the editorial board of ten other journals.

John McCombie is Emeritus Professor of Regional and Applied Economics and Director of the Cambridge Centre for Economic and Public Policy, Department of Land Economy, University of Cambridge, United Kingdom. He is also Senior Emeritus Fellow of the Department, Emeritus Fellow in Economics and Land Economy, Downing College, Cambridge and Director of Studies in Land Economy for Christ's, Downing and Girton Colleges. He is a Fellow of the Academy of the Social Sciences and a Fellow of the Regional Studies Association. He has been a co-editor of *Spatial Economic Analysis* and *Regional Studies*. His research interests and publications include Keynesian economics, the study of national and regional growth rate disparities, economic growth and the balance-of-payments constraint, as well as criticisms of the aggregate production function and conventional measures of the rate of technical progress. He has been a specialist advisor to the House of Lords European Union Sub-committee on the Future of the EU Structural and Cohesion Funds. He has also been an economic consultant to the World Bank and the Asian Development Bank.

Malcolm Sawyer is Emeritus Professor of Economics at the University of Leeds, United Kingdom. He was the lead co-ordinator for the European Union-funded €8 million, 15-partner, five-year project on Financialization, Economy, Society and Sustainable Development (www.fessud.eu). He was managing editor of the *International Review of Applied Economics* for over 30 years, and served on a range of editorial boards. He is the editor of the book series *New Directions in Modern Economics* (Edward Elgar Publishing) and co-edits (with Philip Arestis) the annual series *International Papers in Political Economy* (Palgrave Macmillan). He is the author of 12 books (most recent being *Can the Euro Survive?*, Polity Press, 2017). He has edited or co-edited over 30 books. He has also published over 100 papers in refereed journals on a wide range of topics and recently including papers on financialization, the political economy of the euro area, fiscal policies and alternatives to austerity, alternative monetary policies, and public–private partnerships.

Mario Seccareccia is Professor Emeritus at the Department of Economics, University of Ottawa, Canada, where he taught from 1978 to 2018 in the fields of macroeconomics, monetary theory, labour economics, history of economic thought, and economic history, subjects on which he has also written extensively. He has published some 120 academic articles in

scientific refereed journals or chapters of books, and has authored or edited a dozen books. He has also edited or co-edited over 40 special issues of journals. Many of these publications are of an interdisciplinary nature and cover many areas of political economy. He has been visiting professor in a number of universities in France (University of Bourgogne, University of Grenoble, University Paris 13, and University Paris-Sud) and in Mexico (Universidad Nacional Autónoma de México), and participates regularly in policy debate in both Europe and North America. Among other activities, since 2004 he has been editor of the *International Journal of Political Economy*.

Mark Setterfield is Professor of Economics in the Department of Economics at the New School for Social Research and is also a member of faculty at Eugene Lang College, United States. He is an Associate Member of the Cambridge Centre for Economic and Public Policy at the University of Cambridge, United Kingdom, a Member of the Centre d'Économie de l'Université Paris Nord at the University Sorbonne Paris Nord, France, and a Fellow of the Forum for Macroeconomics and Macroeconomic Policies at the Macroeconomic Policy Institute of the Hans-Böckler Foundation, Germany. He was previously Maloney Family Distinguished Professor of Economics in the Department of Economics at Trinity College, Hartford, Connecticut, United States.

Nat Tharnpanich is Trade Officer (Senior Professional level) at the Ministry of Commerce, Thailand, where he assists the Deputy Prime Minister and Minister of Commerce in Thailand's trade policy-making process. He graduated with a Bachelor degree in Economics (First-class honours) from Thammasat University, Thailand, and later obtained a Master's degree in Economics from the University of North Carolina at Chapel Hill, United States, and a PhD in Land Economy from the University of Cambridge, United Kingdom. His research interest is in the role of non-price competitiveness in promoting economic growth and development. He has published in the *Journal of Post Keynesian Economics* and served as a consultant at the Development Research Group of the World Bank Group.

Jan Toporowski is Professor of Economics and Finance at the School of Oriental and African Studies, University of London, United Kingdom, Visiting Professor of Economics at the University of Bergamo, Italy, and Professor of Economics and Finance at the International University College, Turin, Italy. He studied economics at Birkbeck College, University of London, and the University of Birmingham, United Kingdom. He has worked in fund management, international banking, economic consultancy and central banking, and has published widely on money, finance and economic development, monetary policy and the history of economic thought. He is the author of a two-volume biography of Michał Kalecki.

Acknowledgements

The editors would like to thank all the contributors to this book for their collaboration in preparing this volume to enhance the understanding of economic analysis in a pluralistic perspective. They also wish to express their gratitude to Edward Elgar Publishing for their enthusiastic and professional support during the development of the book. Finally, they are most grateful to Carryl Oberson and Maurizio Solari for their excellent research assistance.

Louis-Philippe Rochon and Sergio Rossi

Introduction: the urgent need for a heterodox approach to economic analysis

Louis-Philippe Rochon and Sergio Rossi

The purpose of studying economics is twofold. First, it is to be able to propose policies that will help to address the various problems of the present. Second, as British economist Joan Robinson (1980, p. 75) once quipped, it is 'to learn how to avoid being deceived by economists'. This is especially true today, given that economics has become mired in overly complicated models and mathematics that more often than not are detached from the reality in which we live, and where the art of telling the story gets lost. The overall purpose of this book is to ensure that students of economics learn how not to be deceived by mainstream economics, while able to also propose realistic policies.

Yet, before being able to propose such policies, the economist has to jump through a few hoops. In fact, there are three such hoops, or stages, that must be complete. But beware: at each step of the way, there are plenty of opportunities for biases, errors and ideological influences.

Stage one consists of the observation of the real world. Here, economists must observe their surroundings, and begin to ask the right questions. What is required, according to British economist John Maynard Keynes (1938, p. 296), is a 'vigilant observation of the actual working of our system'. Yet, this stage is fraught with possible errors and biases. For instance, imagine you are driving down the road and you see workers protesting and picketing. Is your first reaction to honk your horn in support of the workers who are demanding more pay, or is it to yell at them and accuse them of being greedy? The same observation can be interpreted differently, hence the reason we need a 'vigilant' observation of the 'actual workings' of our economy.

Stage two is to set these observations within a proper frame of analysis; in other words, these observations must fit within a greater theory. Yet again, this is not an easy task. For instance, why were so many economists unable to predict the global financial crisis that erupted in 2007–08? The easy and simple answer is that the existence of such a crisis does not fit within mainstream economic theory, which denies the possibility of a crisis. So when one occurred, economists were unable to properly identify it. As a result, while it was becoming clear to many that a crisis was happening, many more denied it, and in fact went to great lengths to claim that 'everything was fine'. Chapter 2 of this book gives ample quotes to show that many economists were unaware of the existence of the crisis (and unable or unwilling to point it out). These economists failed miserably at a 'vigilant observation of the real world'.

Stage three is the empirical step. Indeed, ultimately, everything in economics is an empirical question. If one says, for instance, that reduction in real wages reduces unemployment, the economist must be able to find the proper data and test whether this statement is true (which, in fact, it is not). There are many possible mistakes that can occur at this stage: are you using the correct data; are you applying the correct empirical tests; is your model realistic and representative of the real world?

Finally, the last stage of the economist's work is the policy proposal phase. After having made observations, after having fitted them within the proper theoretical frame, after having tested them with the proper data, the economist can now begin to propose policies that will help to remedy the myriad of problems, such as economic growth, unemployment, income distribution, and more. But this is perhaps the most precarious stage: policies must come from good theory, and a good or 'vigilant' observation of the real world, otherwise wrong policies can have detrimental effects. Two of the worst examples of these are relying on fiscal austerity to stimulate economic activity (which led to the opposite effect), and raising interest rates when the economy is still depressed, which occurs all too often. As Keynes (1936, p. 3) reminds us in *The General Theory*, bad policies stemming from mainstream economics can be 'disastrous if we attempt to apply it to the facts of experience'.

This brings to mind another great quote by Keynes, this time in a publication entitled 'The great slump of 1930', where he notices that 'to-day we have involved ourselves in a colossal muddle, having blundered in the control of a delicate machine, the working of which we do not understand. The result is that our possibilities of wealth may run to waste for a time – perhaps for a

long time' (Keynes, 1930/1978, p. 126). This textbook will make obvious that this 'colossal muddle' is precisely the result of bad policies, bad empirical tests, bad theory and bad observations of the real world.

The purpose of this textbook is twofold. First, it aims at explaining that there is an ever urgent need to abandon mainstream economic analysis; and second, that a credible alternative exists. All authors in this book propose a truly heterodox approach to economic analysis, as regards both macroeconomic theory and economic policy. To be sure, the global financial and economic crisis that erupted in 2007–08 illustrates this need; this crisis is eventually a crisis of economics, since it originates in an essentially wrong approach to the working of our economic systems. Hence, it does not come as a surprise that the majority of economic policy actions taken in the aftermath of this crisis do not work as expected by their proponents. In fact, neither 'fiscal consolidation' (that is, austerity) measures nor 'quantitative easing' policies can live up to their promises, which amount to wishful thinking (to jump-start the economic engine dramatically hit by the crisis). A fundamentally different approach to economic analysis is actually necessary in order to understand and eventually solve this crisis for good.

In this introduction, we provide a detailed overview of the contents of this volume, and point out its distinguishing features with respect to orthodox thinking. We thereby show that another, largely different perspective is required to avoid the fundamental flaws of orthodox economic analyses. This allows us also to point out the need for pluralism in economic research and education, because the lack of it led the economics profession astray under the neoliberal regime that has been increasingly dominating the global economy since the demise of the Bretton Woods system in the early 1970s.

Chapter 1, written by the co-editors of this textbook, explains the meaning and purpose of economic analysis. The authors present and criticize the mainstream definition of economics, which aims at the 'efficient allocation of scarce resources'. This definition includes three main concepts, each with very specific and powerful meanings in economics, namely, efficiency, allocation and scarcity. In this view, the main issue to address thereby is how to allocate a given supply of resources. The chapter also briefly discusses the differences between microeconomics and macroeconomics, emphasizing how mainstream macroeconomics is crucially built on the key assumption of aggregating individuals' behaviour, that is, on microeconomic foundations. Rochon and Rossi also point out that there is no need for money to exist

in mainstream models. Indeed, orthodox models explaining consumption, investment and economic growth contain no money essentially. Money is introduced much later, as part of a discussion about the banking system, as an afterthought, or as an attempt to make these economic models appear more realistic. This is the reason why the first chapter offers an alternative interpretation of the scope and contents of economics.

First, the authors argue that macroeconomics should not be a simple aggregation of individual behaviour and microeconomic magnitudes, that there are characteristics special to macroeconomics, and that social classes (or macro-groups) play an important role in determining economic outcomes. Based on macro-groups, economic dynamics become very important in explaining consumption, investment, prices and economic growth. Further, by emphasizing groups, one can ask a different set of questions and cast these questions within the framework of political economy rather than economics, as clearly explained in Chapter 3. Chapter 1 shows thereby that markets are not free, but are governed by laws and institutions that play a central role in any economic activity. Moreover, in casting this view with regard to social groups, the importance of power becomes paramount: notably, the power over the determination of wages, the power over access to credit and the power of the state. Ultimately, we live in a money-using economy, so, as Schumpeter argued, money should be introduced at the beginning of the discussion of economics (Chapter 4 delves into this subject matter in more detail). That is why this textbook, in contrast to all other macroeconomics textbooks, begins with an explanation of money (Chapter 4) and the banking system and finance (Chapters 5, 6 and 7), after a survey chapter on the history of economic theories (Chapter 3), which is required in order to understand the general framework of any economic analysis, be it theoretical or policy-oriented.

Chapter 2, again by the co-editors, is new to this second edition. In this chapter, the authors argue how mainstream economics is in total disarray. Before the global financial crisis that erupted in 2007–08, there appeared to be some consensus as to what macroeconomics was about, and in particular about which policies mattered. The financial crisis, however, revealed how the economic emperor had no clothes. Indeed, it became evident that mainstream economic models were incapable of explaining the causes of the crisis, and the latter has shown how ineffectual those models are in proposing policies that matter. Indeed, the crisis did more than that. Relying on a number of quotes, the authors of this chapter show how deep the divisions between economists have become, thereby paralysing the profession at large, and revealing the limitations of mainstream economics.

Chapter 3, contributed by Heinrich Bortis, presents the bigger picture within which economic theories have developed, ranging between two camps, namely, economics and political economy. The first section of this chapter provides the essential reason for studying the history of economic thought: dealing with differing or even contradictory theories of value, distribution, employment, and money induces one to independent and open-minded thinking, that is, what John Maynard Keynes (1926) called the 'emancipation of the mind'. This should enable students of the history of economic thought to distil the most plausible theoretical principles, which are the grounds on which policy proposals may be eventually made. Indeed, economic theorizing must be based on the history of economic thinking to have an informed broader picture of the state of the art. In light of this, the second section presents two broad groups of theories – economics and political economy – to bring into the open the fundamental differences in economic theorizing. In economics, the great problems (value and price, distribution and employment) are market issues essentially, and money is neutral. By contrast, the starting point of political economy is the social and circular process of production: the fundamental prices are the prices of production, not market prices; income and wealth distribution are governed by social forces, and employment by effective demand; money and finance play an essential role. The third and fourth sections sketch the historical development of economics and political economy, respectively. Economics starts with Adam Smith, who conceived of the economy and society as a self-regulating system. Jean-Baptiste Say (a follower of Adam Smith) claimed therefore that there can be no unemployment. The great systems of economics were then created in the course of the Marginalist Revolution (1870–90). Léon Walras worked out the general equilibrium model; Alfred Marshall the partial equilibrium approach. Both became constitutive of contemporary mainstream economics. By contrast, the French surgeon François Quesnay is at the origin of the political economy line. He considered the flows of goods and money within the social and circular process of production to produce the net output at the free disposal of society. As regards production, David Ricardo worked out the labour value principle and the surplus principle of distribution. Piero Sraffa revived the classical (Ricardian–Marxian) approach, which had been submerged by the Marginalist Revolution. John Maynard Keynes elaborated the principle of effective demand, represented by the multiplier relation, implying the existence of involuntary unemployment. At the time of writing, the post-Keynesian and classical–Keynesian followers of Sraffa and Keynes, together with Marxists, form the core of modern political economy, representing an alternative to the neoclassical mainstream. The last section of Chapter 3 discusses the plausibility of these two approaches. The capital-theoretic

debate emerges thereby as the theoretical watershed between economics and political economy.

Chapter 4 focuses on monetary economies of production. Everybody knows the old song, 'Money makes the world go round'. In reality, this is a good approximation of how our economic system operates. Indeed, as Louis-Philippe Rochon explains in this chapter, we live in a monetary economy. This means that we cannot purchase goods without money; we cannot invest without money; we cannot hire workers without paying them a wage in money. Money is indeed at the core of our economic system. While this may be quite apparent to many, money does not feature in neoclassical economics; or if it does, it is merely there to give some semblance of reality to an otherwise unrealistic view of the world. This chapter first discusses the barter view of money in neoclassical analysis, and how in this view what serves as 'money' has evolved through time. In this framework money is introduced to facilitate trade and has no other purposes. In this sense, there is no need for money in discussing employment, wages, supply and demand, investment and economic growth. In fact, there is no need for money even to discuss prices in neoclassical analysis. This chapter offers a criticism and an alternative view, which is focused on the creation and circulation of money, and its relationship with debt, which characterizes a 'revolutionary' approach. To be sure, money is necessary to explain production and employment as well as economic growth.

Chapter 5, written by Marc Lavoie and Mario Seccareccia, elaborates on this. It provides a brief analysis of the historical evolution of money and recalls some debates about it. Since money, in its essence, is merely the outcome of a balance sheet operation, banks play a key role in any monetary economy. The purpose of this chapter is to describe why all aspects of macroeconomic analysis in a modern economy must necessarily involve the monetary system. Monetary relations result from the existence of a group of key institutions in a monetary economy, namely banks; which, together with the central bank, are crucial in the modern payments system and are the purveyors of liquidity to the whole economy. This chapter starts therefore with an explanation of how banks are the principal creators of money in nearly all modern economies, and why, by their very nature, they are private–public partnerships, especially evident at times of crisis when the public 'trust' that is so critical to their existence is broken, therefore requiring a regulatory framework within which the activity of creators of money is severely circumscribed. Lavoie and Seccareccia consider the composition of this creation of money by the banking system, including the central bank, and show why it varies with the performance of the economic system. This is followed by an investigation of the logic of money

creation in the traditional analysis of the monetary circuit and how this has been transformed somewhat under the so-called regime of financialization (which is discussed in Chapter 18), especially with the perverse incentives generated by off-balance sheet operations via securitization. As in most first-year textbooks, Lavoie and Seccareccia begin by discussing the specifics of bank balance sheet operations and how bank money is created endogenously either to finance productive activity, as in the traditional circuit model of financing production, or more recently through the financing of household consumption spending. They thus consider how different forms of spending behaviour by either the private or the public sector lead to the creation or destruction of money. They thereby comment on the role that the public sector plays on the asset side of banks' balance sheets, which is also critical to how the banking sector's net worth is usually re-established after recessions. Finally, this chapter provides a discussion of payments and settlement systems and of the role played by the interbank market, where the central bank can set the overnight rate of interest and thus largely control interest rates in the economy. The interbank rate of interest is a critical instrument for the conduct of monetary policy. This discussion establishes a bridge with Chapters 6 and 7 (on the financial system and monetary policy, respectively).

Chapter 6, by Jan Toporowski, explains notably that the financial system emerges out of the financing needs of production and exchange in a capitalist economy. A special role is played by the financing needs of the state, for which are created the institutional foundations of the long-term debt markets that characterize the modern capitalist economy (that is, stock markets, insurance and investment funds). These markets develop further with financial innovations that provide new scope for financial operations alongside production and exchange. Financial operations are a distinguishing feature of Marxist political economy – in the case of Hilferding's *Finanzkapital* (Hilferding, 1910/1981) and the financial theories of Kalecki and Steindl derived from Hilferding's work – as well as Keynes's macroeconomics and post-Keynesian theories. By way of contrast, consistent with the irrelevance of money in neoclassical economics as explained in Chapter 4, mainstream macroeconomics and portfolio theory do not integrate the financial sector within a theory of how a capitalist economy operates with a complex financial system. The mainstream view also confuses saving with credit, exemplified in the theory of interest. The financial approach to macroeconomic theory is developed, by way of contrast, in the 'financial instability hypothesis' of Hyman Minsky, considering how debt changes over a business cycle. Minsky's view is distinguished from post-Keynesian theories by Minsky's denial of Keynes's interest rate theory of investment, but has contributed the notion of usury (due to excessive debt) common in post-Keynesian theories of financialization.

Chapter 7, written by Louis-Philippe Rochon and Sergio Rossi, points out the specific role of the central bank in domestic payments and settlement systems: as money and credit provider. It focuses on the emission of central bank money as the means of final payment for every transaction on the interbank market. It thereby distinguishes between the monetary intermediation carried out by the central bank as a matter of routine, and necessity; and the financial intermediation that it carries out as a lender of last resort, when its counterparties are not in a position to obtain enough credit on the interbank market. The authors show thereby that the central bank is crucial for financial stability, thus introducing the reader to the need to go much beyond price stability (particularly as measured on the goods market) for monetary policy-making. This chapter then focuses on monetary policy strategies, instruments and transmission mechanisms. Two major strategies (monetary targeting and inflation targeting) are discussed, criticizing their conceptual framework, and observing their macroeconomic costs as measured by so-called 'sacrifice ratios' with respect to output and employment losses. As the authors explain, monetary policy must contribute to macroeconomic stabilization and not just worry about price stability on the goods market. This discussion is elaborated on to present traditional as well as 'unconventional' monetary policy tools, critically considering their consequences for the whole economy. This framework is further expanded to present the transmission channels of monetary policy, considering also the ongoing discussion about regulatory capital and financial reforms as well as banking supervision at national and international levels that aim to influence aggregate demand and thus achieve the relevant monetary policy goals.

Chapter 8, by Stavros A. Drakopoulos, is another new chapter in this second edition. This chapter explains the role of consumption expenditures in modern economies and their significance for the determination of the level of output and employment, both of which are important for the well-being of the population as well as for the orderly working of the economic system. After presenting the theory of intertemporal choice that forms the basis of mainstream consumption functions, the author discusses Keynes's approach to consumption, and particularly his criticism of the standard model of consumer behaviour, his emphasis on the role of consumption for the level of employment, and his analysis of aggregate consumption patterns. As a matter of fact, consumption represents a large part of expenditures on the market for produced goods and services. It is thus important for aggregate demand – that is, the total demand for all goods and services in the economy – since, according to Keynesian theory, aggregate demand determines the level of output and employment in an economy. Also, income that is not consumed is saved, and savings have a large impact on economic growth.

As a result, consumption is important to understand savings, capital stock, investment, employment and income growth. Now, as the effectiveness of economic policy is closely related to the nature of the consumption function, Drakopoulos presents the main mainstream theories of consumption (namely, the life-cycle income hypothesis, the permanent income hypothesis and the random walk theory of consumption), and explores the heterodox approaches to consumption, focusing mainly on the relative income hypothesis. The author is thereby in a position to show the consequences of consumption theories for the effectiveness of economic policies to address unemployment and economic downturns.

Another new chapter in this second edition, Chapter 9 by Thomas Dallery, describes and discusses the determinants of investment. This is an important question, because investment is a driver of business cycles, as it affects economic dynamics. It is through investment that firms prepare the future of production, implement innovations and compete with one another. Dallery presents first the mainstream view of investment, which rests upon the role of price factors and the principle of substitution between production factors. He then explains Keynes's view of investment, emphasizing notably the role of uncertainty, expectations and effective demand. Investment is indeed an ambivalent variable, because it affects both demand and supply. On the one hand, investment increases potential output, and thus contributes to enhance or to improve the supply of goods and services. On the other hand, investment is a source of demand for those firms producing capital goods. Dallery explains that the interplay between these components of investment can induce economic instability. This chapter also provides a synthesis of various heterodox theories of investment, where profitability, demand and financing constraints affect firms' decisions to invest. Last but not least, Dallery gives a hint of the empirical evaluation of the factors determining investment, especially as far as financialization is concerned. Indeed, a number of economists blame financialization for being a cause of declining investment, since at least the eruption of the subprime crisis (see Chapter 18). This decline of investment is all the more worrying in that our economies have plenty of sectors requiring investment (such as for ecological transition, elderly care and education). As Dallery points out, investment is an expenditure made today in order to prepare for the future. With low levels of investment, our society is no longer preparing for its own future. With firms reluctant to invest, it may be a great opportunity for the state to assume its role in socializing investment in these sectors, thereby realizing what Keynes imagined in the 1930s.

Chapter 10, contributed by Jesper Jespersen, expands on the previous chapter and explains that aggregate demand comprises private consumption,

private investment, government expenditure and net exports. In this connection, it points out that neoclassical economists consider aggregate demand as rather unimportant. They argue that total output (gross domestic product, GDP) is determined mainly by the supply of labour and capital, quite independently of demand. They consider the market system as self-adjusting, which leads them to conclude that, in the long run, 'the supply of goods creates their demand'. By contrast, according to heterodox economists and Keynesian macroeconomic theory, aggregate demand is an important analytical concept: it is the major driving force behind the level of output and employment in both the short and the long run. This consideration makes demand management policies instrumental for creating macroeconomic stability and economic growth. Richard Kahn was notably one of the first Keynesian economists who contributed to the theory of aggregate demand: he invented the analytical concept of the investment multiplier as a short-run dynamic phenomenon. This chapter expands on this, presenting Keynesian demand management policies, notably in periods where aggregate demand is lower than potential GDP: fiscal policy, monetary policy and exchange rate policy should be expansionary in recessions, and might be restrictive in boom periods. Demand management policies can thus close output gaps and make the macroeconomic system more stable. Against this background, neoclassical scepticism builds on the assumption that the macroeconomic system is self-adjusting: if wages are fully flexible, then the labour market will adjust by itself to full employment. In such cases aggregate demand, also by itself, adjusts to potential output (GDP) via changes in real wealth and/or changes in net foreign trade. Hence, demand management policies are at best superfluous in the orthodox view. As shown in this chapter, these different analytical outcomes depend on the theoretical macroeconomic framework: Keynesian aggregate demand analysis leaves the future open, which creates room for demand management policy; by way of contrast, the neoclassical assumption of automatic market adjustment makes aggregate demand equal to potential supply of output, thereby excluding both inflation and unemployment over the long run, unless the state intervenes, disturbing this 'free market equilibrium'.

Chapter 11, written by Alvaro Cencini and Sergio Rossi, focuses therefore on inflation and unemployment. It starts with an overview of both pathologies, explaining several basic concepts such as the consumer price index, the purchasing power of money, inflation, deflation, unemployment and stagflation. It then points out their relevance and effects: income redistribution and monetary instability caused by inflation; social and economic disruptions caused by unemployment. The next stage is devoted to a critical analysis of the way mainstream economics has attempted to explain inflation and

unemployment. In particular this chapter shows that inflation should not be confused with a rise in the cost of living and that neither demand-pull nor supply-push price variations cause a loss in the purchasing power of money, as only the agents' purchasing power is affected by a price variation originating in either demand or supply on the market for produced goods and services. It also shows that agents' forms of behaviour cannot reduce total demand and thereby be at the origin of an excess in total supply leading to unemployment. In the last section of this chapter the authors explain that inflation and unemployment are the twin outcomes of one single macroeconomic cause: that is, a pathological process of capital accumulation and overaccumulation. Their analysis is entirely macroeconomic and rests on Keynes's identities between total supply and total demand, and between savings and investment at the macroeconomic level. These two identities necessitate an analysis capable of reconciling them with the numerical disequilibria defining inflation and deflation. With regard to this, the authors argue that from a macroeconomic point of view only pathological unemployment matters and that involuntary unemployment, in Keynes's terms, is the consequence of an excess in total supply, that is to say, deflation. By investigating the process of capital accumulation the authors reveal the mechanisms leading to both inflation and deflation and so make sense of the global economic crisis that erupted after the bursting of the financial bubble in 2007–08.

Chapter 12, by Malcolm Sawyer, focuses on the role of fiscal policy. It explains that fiscal policy relates to the balance between government expenditures and tax revenues, which is important for the level of employment and economic activity. In doing so, the chapter points out that the need for fiscal policy arises from the perspective, generally denied by mainstream economists, that the capitalist economy is subject to cyclical fluctuations as well as to unemployment arising from inadequacy of aggregate demand, as discussed in Chapter 10. As a matter of fact, the private sector exhibits instabilities and suffers from insufficient aggregate demand to underpin full employment and capacity utilization. Fiscal policy is one instrument (amongst a number) to address these features, through public budget positions that offset inadequate private demand, and through automatic stabilizers and discretionary fiscal policy to offset variations in demand. This chapter sets up the arguments relating to fiscal policy in the framework of a closed economy (for simplicity). It explains the equilibrium condition that (in *ex ante* terms) injections equal leakages: $I + G = S + T$ (with I private investment, G government expenditure, S private savings, and T tax revenues). If $S > I$, then $G > T$ is required. The chapter explains the implications of that in terms of realization of savings and for public deficits. As regards the mainstream view, the equilibrium $S = I$, and the mechanisms by which this is said to occur, refer

to the ('natural') interest rate. Further, in this view, the so-called 'Ricardian equivalence' between agents' taxation and public debt to finance government spending leads to the 'crowding out' effect of expansionary fiscal policies, as explained in Chapter 10. By contrast, the heterodox view that I and S tend not to be equal explains that public deficits can be funded when $S > I$. Indeed, appropriate public deficits do not put pressure on interest rates, particularly in a recession, because there are excessive savings with respect to desired investment by the business sector of the economic system. This chapter also briefly discusses the so-called 'functional finance' approach, thus pointing out that the objective of fiscal policy should be high levels of employment and not balanced budgets *per se*.

In Chapter 13, Mark Setterfield focuses on economic growth and development. He first provides a definition of economic growth, followed by an overview of its statistical record. Three salient features of the growth record are thus emphasized: the extent and unevenness of economic growth (resulting in the processes of 'catching up' and – for most – 'falling behind'); the unbalanced nature of economic growth (resulting in structural change, as exemplified by deindustrialization); and fluctuations in the pace of economic growth over time (which can be interpreted either as growth cycles or as discrete and historically specific episodes of growth). This chapter then considers whether economic growth is a supply-led or demand-led process, contrasting mainstream neoclassical (supply-led) and alternative post-Keynesian (demand-led) views. Simple analytics (such as shifting production possibility frontiers, with economies operating either on these shifting frontiers – neoclassical case; or within their interiors – post-Keynesian case) are used to illustrate the differences between views. Next, the chapter invites further consideration of the post-Keynesian (demand-led) view of economic growth. Both Kaleckian and Kaldorian growth theories are thus sketched using simple analytics (on a par with multiplier analysis), drawing attention to the different sources of autonomous demand that each theory considers to be the key driver of long-run growth. Various properties of economic growth (inspired by the post-Keynesian view of the growth process) are then considered. These include the paradox of thrift (an increase in the saving rate is harmful to economic growth); income distribution and economic growth, including discussion of the fact that raising the profit share of income may be harmful to GDP growth; technical progress ('Verdoorn's Law'); and interaction of supply and demand in long-run growth (endogeneity of potential output to actual output; the two-way interaction of actual and potential output, resulting in self-reinforcing virtuous and vicious circles of growth). This last topic invites further consideration of the role of supply factors in demand-led growth theory, as taken up in Chapter 20 (on sustainable development).

Chapter 14, by Omar Hamouda, expands on income and wealth distribution. Whether from the micro- or the macroeconomic perspective, the purpose of studying economics has always been and will continue to be focused on two fundamental questions: (1) How and what to produce in terms of goods and services? (2) Who receives what from the ongoing creation of wealth? Whether the economic aspects of these two questions are interrelated and uniquely determined, as the neoclassical approach maintains, or are determined separately by different sets of forces, as held by many other schools of thought, the implication – in terms of understanding what is meant by income distribution – is subject to controversial interpretations. The object of this chapter is to expose and elucidate various perspectives on the concepts of income share, and income distribution and redistribution in relation to the distribution of contribution and effort. In this chapter, income distribution is understood as the share that the remunerating factors of production get for their contribution to the creation of wealth. Income redistribution refers to income reallocation among members of a community regardless of whether or not they participate in the creation of wealth. In different eras in history, stages of economic development, political structures as well as the type of moral guidance dictated how, when and who gets what from the proceeds derived from the common effort in producing what is required to satisfy human needs. This chapter is thus devoted first to exploring the general conception of shared income derived from the wealth of a nation. This wealth is referent to more than just economic criteria, especially in pre-industrial economies or those that do not adhere to market forces. The focus of this chapter then turns to how income is distributed under purely economic considerations, as explained and justified in the classical and the neoclassical schools of thought. Finally, the chapter reviews how the post-World War II 'welfare state' has slowly shaped and created a system of income redistribution decided by economic policies and regulation.

In Chapter 15, Robert A. Blecker focuses on trade and development. He begins this chapter with a brief review of the conventional arguments for free trade in general, and for a policy of economic openness as the best route to long-run growth and development. In fact, exports can be an important part of a development strategy, especially when they grow as part of a virtuous circle of industrialization and rising incomes at home, but policies of pure free trade have not historically been the most successful. The most successful developing nations since the late nineteenth century have been those that have combined a significant role for the public sector in economic management with active promotion of exports and selective reliance on markets (but not complete state domination of markets). In addition, the comparative advantage model rests on the twin assumptions of full employment and

balanced trade, which do not generally hold in reality, and without which the global trading system does not necessarily work in the mutually beneficial ways implied by standard theories. In the real world, some countries use mercantilist policies (which this chapter defines) to foster absolute competitive advantages and to increase national employment at the expense of their trading partners. As a result, mercantilist strategies have generally outperformed neoliberal strategies based on open markets with little state direction. The chapter also includes a brief discussion and update of Joan Robinson's view of 'new mercantilism', for example as regards currency undervaluation (more recently dubbed 'currency war'). Multinational corporations and global financial investors can profit from free access to foreign markets and foreign mercantilist practices (such as undervalued currencies and repressed wages), leading to a world in which outsourcing or offshoring undermines distributional equity and the social fabric in the advanced economies without guaranteeing equitable development in the global South.

These issues are further elaborated upon in Chapter 16 by John McCombie and Nat Tharnpanich, who focus on balance-of-payments constrained growth. Neoclassical growth theory explains differences in economic growth from the supply side within the framework of a closed economy. Economic growth is thus modelled with respect to an aggregate production function, with either exogenous or endogenous technical change, and the economy is assumed to be at full employment. However, it is clear that in many cases the economic growth rates of countries are interlinked through trade flows and the balance of payments. The concept of a balance-of-payments equilibrium growth rate refers to the growth rate consistent with equilibrium on the current account, or where there is a sustainable growth of net capital inflows. If this growth rate is below the growth of productive potential, the country is said to be balance-of-payments constrained. In these circumstances, it will have a lower rate of induced technical change, a lower rate of capital accumulation, increasing inefficiency and greater disguised unemployment. This chapter outlines this approach. It shows that, in the long run, the economic growth of a country is determined by the growth rate of its exports (operating through both the Harrod foreign trade multiplier and the Hicks super-multiplier). The key to the growth of exports is primarily a country's degree of non-price competitiveness, which is determined by the value of its income elasticity of demand for exports and imports relative to that of other countries. In this model, unlike the neoclassical growth model, economic growth is demand-driven. This chapter shows how the economic performance of one country, or group of countries, can constrain the economic performance of another country or group. It also shows how economic growth of a country (y) may be explained by the simple rule $y = x/\pi = \varepsilon z/\pi$, where x

is the growth rate of exports, π is the domestic income elasticity of demand, ε is the income elasticity of demand for exports, and z is the growth rate of the country's export markets. While the full model is more complicated than this, this simple rule, known as 'Thirlwall's Law', provides a good empirical explanation as to why economic growth rates differ around the world.

Sergio Rossi, in Chapter 17, focuses on European monetary union, presenting its own history and the workings of its main institutions since the 1960s. He also explains the euro area crisis, pointing out its monetary and structural factors. He thereby shows the fundamental flaws of the single-currency area as well as of its anti-crisis policies at both national and European levels. The discussion is expanded to present an alternative path to European monetary integration, in the spirit of Keynes's International Clearing Union based on a supranational currency unit, which in fact can be issued without the need for its member countries to dispose of their monetary sovereignty and thereby preserves national interest rate policies as a relevant instrument to steer economic performance at the euro area level. The chapter sets off from the original proposals for European monetary union, as put to the fore in the Werner Plans of the early 1970s, as a halfway station between so-called 'economists' and 'monetarists'; the former (led notably by Germany) being in favour of a convergence of macroeconomic magnitudes before European countries may actually enter into a fixed exchange rate regime, and the latter (including France and Italy) favouring an early fixing of exchange rates, which they considered as a factor of economic convergence for the countries involved. The chapter focuses then on the Delors Plan, which led to the 1992 Maastricht Treaty and the adoption of the European single currency in 1999. It further explains the institution and workings of the European Central Bank, including its governance and lack of accountability with regard to its (single) monetary policy goal. The next section expands on this, to explain that the euro area crisis does not really originate in 'excessive public deficits' and debt as a percentage of GDP in the Maastricht sense. As a matter of fact, it is a monetary–structural crisis originating in the institutional design and workings of the European single currency area. Beyond increasing intra-euro area trade imbalances and speculative capital flows, the European single currency has given rise to payment imbalances – as captured by the TARGET2 payments system – which need a symmetric rebalancing to avoid the depressionary spiral induced by widespread fiscal austerity and monetary inactivity by the European Central Bank. The chapter therefore presents an alternative path to European monetary integration, which is akin to 'Keynes's Plan' presented at the 1944 Bretton Woods conference, since it reintroduces national currencies while making sure that all intra-euro area international payments are finalized between the relevant national central banks through the emis-

sion of a scriptural means of final payment by an international settlement institution resident at the European Central Bank, the Bank for International Settlements or the International Monetary Fund. The discussion of this alternative refers also to Robert Triffin, notably to his own critiques of a single currency area for Europe and to his alternative proposal in that regard.

Chapter 18, by Gerald A. Epstein, addresses financialization, that is, the growing and excessive importance of financial motives, markets and institutions in the working of today's capitalist economies. According to standard macroeconomic analysis, finance in a capitalist economy serves households, non-financial businesses and governments in several ways (as explained in Chapter 6): by providing households with safe places to store their savings, and by channelling these savings to productive and profitable uses by non-financial corporations; by providing opportunities for households to save for their retirement; by allowing corporations to mobilize large sums of capital for investment in productive enterprises; and by providing ways for households and firms to insure against risks. According to this idealized view of the world, there is a clear separation between the financial sector and the 'real sector', and finance prospers when it serves the 'real economy'. In contrast to this idealized vision, modern capitalist economies seem to be structured in a starkly different way. As shown by the global financial crisis that erupted in 2007–08, the financial sector in many advanced capitalist economies has been operating in a rather closed loop, in which it enriches itself, often at the expense of households and non-financial corporations, rather than serving the 'real economy'. Further, the chief executive officers and boards of directors of non-financial corporations manage many of these corporations as if they were simply portfolios of financial assets to be manipulated to maximize the short-term income of small groups of corporate executives, rather than the long-run growth and profitability of corporations. In short, there is increasing evidence that many modern capitalist economies have become 'financialized', and this 'financialization' has contributed to significant economic and social ills, including huge financial crises, increased inequality between the top managers of corporations and everyone else, and to stunted investment in long-term productivity growth. Epstein presents the most important theories of financialization, gives a brief historical description of the evolution of financialization in contemporary economies, and surveys the key empirical evidence on the impacts of financialization on instability, inequality and economic growth. The author explores many dimensions of financialization, namely its impact on financial instability, its impact on investment and employment decisions by non-financial corporations, its impact on public finance and public deficits, the role of financialization in generating inequality, and public policies for confronting the problems asso-

ciated with financialization, such as financial transactions taxes, restrictions on stock options, and industrial policy.

In Chapter 19, Robert Guttmann expands on this to explain economic imbalances and crises. There is overpowering evidence that our capitalist system is subject to endemic imbalances, which, if large and/or persistent enough, lead to crises. This is a recurrent pattern in which underlying imbalances and crises enter into a dialectical relationship, with crises serving as adjustment processes that may (or may not) resolve the imbalances triggering them in the first place. Whereas orthodox economic theory tends to treat crises as exogenous shocks intruding from the outside to upset our supposedly self-balancing system, we need to understand this phenomenon instead as intrinsic to capitalist economies. The key to this reinterpretation effort is to pinpoint the imbalances that such a system gives rise to as a matter of its normal *modus operandi*. For a long time, economists of all stripes have analysed imbalances between demand and supply at the micro level of individual actors or the sectoral level of markets and industries, primarily to show that affected parties respond to any disequilibria in such a way as to eliminate any excess of demand or supply. Their argument, crystallized in the famous Marshallian cross of intersecting demand and supply schedules, works with appropriate changes in the price level, sending corresponding signals for both sides of the marketplace to respond to. But when structural changes in our economic system led to increasingly pre-set prices that resisted falling in situations of excess supply, the standard market equilibrium argument broke down. In the face of growing downward price rigidity, suppliers would end up slashing output and employment rather than prices. It was up to John Maynard Keynes to identify, in the midst of the Great Depression, the horrifying fallacy of composition, where what was good for individual actors was disastrous for the system as a whole, as expenditure cutbacks of some would impose income losses on others and so trigger additional cuts in spending. Keynes's solution, namely to bring in the government as an extra-market actor not bound by private sector budget constraints and profit motives to boost total demand in the economy, gave rise to a revolution in economic thinking and transformed our system into a mixed (private–public) economy subject to active crisis management by the state. Keynes's emphasis on inadequate demand generated by the private sector represents only one side in an age-old debate among that minority of economists seeking to explain the cyclical fluctuations of our economy's growth pattern. Juxtaposing Keynes's underconsumption argument has been Marx's emphasis on overproduction according to whom the capitalists' incessant chase for greater profits ('surplus value'), motivated by their competition with each other, would inherently drive supplies beyond the system's limited absorption capacity.

The underconsumption versus overproduction argument is given added weight when one links (cyclical) economic growth patterns to (functional) income distribution, as has been attempted by Kalecki. The Kaleckian link between economic growth and income distribution integrates micro-level actions (mark-up pricing, investment function) and macro-level determinants (wage and profit shares) to yield a more profound insight into capitalism's inherent demand–supply imbalance. It points directly to what unites all three of these heterodox masters (Marx, Keynes and Kalecki), which is grounded in the relationship between wages and productivity. As crystallized in Keynes's notion of 'efficiency wages', the balanced growth of our economic system depends on both of these variables growing at pretty much the same rate. Some sort of crisis will occur when productivity levels outgrow wages or vice versa. Business cycle theory has also focused on the role of credit, a major factor in the destabilization of our economy's growth pattern as recognized by Mikhail Tugan-Baranovsky, the Austrians (such as Friedrich von Hayek), and above all by Hyman Minsky, whose 'financial instability hypothesis' has made it irrefutably clear that business cycles are to a significant extent driven by parallel credit cycles. In the Minskian world, the inevitable accumulation of excessive debt renders economic actors increasingly vulnerable and eventually forced to deleverage when acute incidences of financial instability arise to trigger recessionary adjustments. His emphasis on the relationship between debt servicing charges and income levels identifies, besides demand–supply and wage–productivity imbalances, a third crucial imbalance capable of triggering a crisis. Minsky highlighted another aspect of crucial importance: the presence of long waves around which our (relatively short-term) business cycles are woven. While he stressed the financial factors underpinning such long waves (implying a 'financial supercycle' of leverage-enhancing financial innovations and growing propensity for risk), others have approached the subject of these longer-term phases of rapid and slow growth as a matter of commodity price movements (Nikolai Kondratiev), bursts of technological change (Joseph Alois Schumpeter), or institutional transformations (the French Regulation School). One major insight from long-wave theory is that some crises, especially those occurring either at the beginning or towards the end of the wave's downward phase, are more serious – in both depth and length – than normal cyclical downturns. Such 'structural' crises deserve special attention. Their imbalances typically engulf the entire system, cannot be resolved by crisis as adjustment process, and need reform. In today's globalized economy, and as sharply confirmed by the unfolding of the latest structural crisis from late 2007 onward, we have a new source of crisis-prone imbalance to consider: in the absence of corrective adjustment mechanisms, a major flaw in how globalization has played out, the interaction dynamics of chronic deficit and surplus countries creates

tensions within the world economy that threaten to explode in new types of crises: currency crises, sovereign debt crises, ecological crises or resource supply crises. The story of capitalism is one of transformational imbalances and many-faceted crises; a fact that economists yet need to acknowledge.

In Chapter 20, focusing on sustainable development, Richard P.F. Holt lays the foundations for a sustainable approach to economic development, which incorporates quality of life and sustainability in ways that the neoclassical model does not, by looking at different capital stocks and distinguishing between development and economic growth. The traditional assumption that more economic growth means a higher quality of life is being questioned around the world. Economic development as defined in this chapter means a broad-based increase in the standard of living, which includes quality of life and sustainability, while economic growth is described as any increase in undifferentiated output or income. Economists and policy-makers have often used the terms interchangeably. This chapter makes a clear distinction. Besides scholarly work that makes this distinction, there are also popular and political groups advocating for such a distinction. For example, efforts by the United Nations through the United Nations Human Development Index and alternative domestic measures such as the 'genuine progress indicator' (GPI) are based on the capabilities and sustainability of development that improves the quality of life, as well as sustainability, that go beyond traditional definitions and measurements of economic growth. Many communities in both industrialized and developing countries are developing locally based indicators of sustainability or quality of life to supplement traditional economic measures. All of these measures show an increasing recognition that economic development depends on more than raising national income. This is based in part on the reality that economic prosperity depends on environmental and social sustainability in ways that the neoclassical model has not addressed adequately. It also reflects a desire for balancing economic well-being with other aspects of well-being such as health and human relationships. The discussion in this chapter goes beyond traditional views. It focuses on both positive and negative impacts of economic growth on quality of life, sustainability, and how the fruits of this growth are shared among income groups important for economic prosperity. The term 'sustainability' is used most often in discussions of natural resource depletion or carrying capacity of the environment. This chapter uses it in its broader framework to include all inputs necessary to maintain a given standard of living or quality of life. This includes what economists call 'human capital', such as skills and health of a population, and the 'social capital' of viable private and public institutions. This chapter also argues that thinking and acting locally as well as globally must address concerns about economic development. This is not to

argue that national and international policy or economic performance have no effect on local conditions. They have important consequences for communities and the individuals who live in them. But economic, geographic and social realities cause substantial regional variation in economic growth and development in all countries.

The last chapter, by John King, offers a conclusion to this textbook, raising the question of the need of microfoundations to macroeconomics. The chapter begins by asking why we might need a separate (sub)discipline of macroeconomics, relatively autonomous from microeconomics. The author uses the example of unemployment: excess supply cannot be eliminated by reducing prices in a macroeconomic context, as Keynes (1936) rightly explained in Chapter 19 of *The General Theory*. This is a fallacy of composition. Yet, as King notes, supporters rarely set out the mainstream case for microfoundations, as it is simply taken for granted. As King explains, microfoundations are an example of question-begging or persuasive language, which is often found in economics. Consider, for example, 'free market economics' (in its Australian translation, 'economic rationalism'). Who would ever want to be unfree (or irrational)? 'Microfoundations' is also a metaphor. King notably provides examples of other metaphors that are used in economics, and discusses thereby the criteria for identifying good (and bad) metaphors. He thus suggests that 'microfoundations' is a bad metaphor, which would have the consequence – if generally accepted – of destroying macroeconomics as a separate, relatively autonomous (sub)discipline. Insistence on providing microfoundations can also be seen as part of the 'economics imperialism' project, since if it were applied to the other social sciences it would also destroy their autonomy. Further, the case against microfoundations can be reinforced by considering a number of additional fallacies of composition, which are sometimes described as paradoxes. The most familiar is the 'paradox of thrift'. The global financial crisis has brought to prominence the 'paradox of liquidity' (or deleveraging). These are considered in some detail in this chapter, where the author also devotes some space to the (Kaleckian) 'paradox of costs', and the potential for wage-led economic growth to restore prosperity, for instance in the euro area as a whole and in a number of its member countries in particular. This takes us back to John Maynard Keynes and the absence of any unambiguous macroeconomic connection between wages and unemployment. Further, what Paul Davidson (2003–04) criticized as 'imperfectionism' is certainly wrong; our paradoxes would still apply, even if all markets were perfectly competitive and all prices (and wages) were perfectly flexible downwards. Hence, removing market imperfections would not eliminate the fallacy of composition or reduce the importance of downward causation. To conclude: language matters in economics. So

does methodology. Students of economics should be prepared to learn from other social sciences, and from the philosophy of science. And so should their teachers.

REFERENCES

Davidson, P. (2003–04), 'Setting the record straight on *A History of Post Keynesian Economics*', *Journal of Post Keynesian Economics*, 26 (2), 245–72.

Hilferding, K. (1910/1981), *Das Finanzkapital. Eine Studie über die jüngste Entwicklung des Kapitalismus* [*Finance Capital. A Study of the Latest Phase of Capitalist Development*], London: Routledge & Kegan Paul.

Keynes, J.M. (1926), *The End of Laissez-Faire*, London: Hogarth Press.

Keynes, J.M. (1930/1978), 'The great slump of 1930', in E. Johnson and D. Moggridge (eds), *The Collected Writings of John Maynard Keynes, Vol. IX: Essays in Persuasion*, London: Royal Economic Society, pp. 126–34.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Keynes, J.M. (1938), 'Letter 1267, 4 July 1938', in E. Johnson and D. Moggridge (eds), *The Collected Writings of John Maynard Keynes, Vol. XIV: The General Theory and After. Part II, Defense and Development*, London: Royal Economic Society, pp. 296–7.

Robinson, J. (1980), *Collected Economic Papers, Volume 2*, Cambridge, MA: MIT Press.

Part I

Economics, economic analysis and economic systems

1

What is economics?

Louis-Philippe Rochon and Sergio Rossi

OVERVIEW

This chapter:

- critically discusses the mainstream definition of economics, based on the 'efficient allocation of scarce resources';
- explains the differences between microeconomics and macroeconomics, and emphasizes how mainstream macroeconomics is built on the assumption of aggregating individuals' behaviour, that is, on microeconomic foundations;
- shows that neoclassical models explaining consumption, investment and economic growth contain no money essentially, thereby being fundamentally inappropriate to understand the real world;
- argues that macroeconomics should not be a simple aggregation of individual behaviour and microeconomic magnitudes, and that social classes play an important role in determining economic outcomes.

This analysis explains thereby that markets are not really free, but are governed by laws and institutions, which play a central role in any economic activity. Further, in casting this view in terms of social groups, the importance of power becomes paramount: power over the determination of wages, power over access to credit and the power of the state; all elements that have been missing and cannot be introduced in orthodox economics.

Introduction

In his 1924 tribute to Alfred Marshall, his former teacher, British economist John Maynard Keynes (1924, p. 322) noted that:

> the master-economist must possess a rare combination of gifts. He must be mathematician, historian, statesman, philosopher – in some degree. He must

> understand symbols and speak in words. He must contemplate the particular in terms of the general and touch abstract and concrete in the same flight of thought. He must study the present in the light of the past for the purposes of the future.

Keynes's insights were profound, and they place a considerable burden of knowledge on economists. This is no surprise, as the stakes are very high. After all, there are a number of economic issues that need to be addressed and explained, and economists play a central role in explaining them, from the wages we get paid, to the jobs we have or the jobs so many of us would like to have, to the interest rates we pay on our credit cards, bank loans and mortgages, to taxes and the prices of goods and services.

Economic policy also plays an important role. Should governments involve themselves in the workings of markets, and if so, what should this involvement be? What role should central banks play? While people are happy when the economy is growing soundly, many lose their jobs and even their homes when the economy goes bad. Economics impacts us all.

An economist must therefore be able to understand the world around them and to propose solutions to the problems we face. The list of problems is a long one: unemployment, inflation, income inequality between the rich and the poor, pollution of the environment, economic growth and recessions, and many more.

To understand the world, the economist must possess knowledge that goes well beyond the strict confines of 'economics'. What Keynes was suggesting in the above quote is that economics is not a science that can be studied in isolation, at the same level as physics or chemistry, but it borrows many elements from other disciplines, especially from other social sciences. Indeed, to be a good economist – and this has never been truer than today – economists must adopt a pluralist approach, and recognize that political science, history, psychology, sociology, mathematics, logic, ecology and philosophy, to name but a few, provide indispensable inputs in forming the economist's mind. In other words, if the goal of the economist is to better understand the world we live in, or what we call the 'real world', rather than some hypothetical world, they must keep an open mind and build bridges with other disciplines. To ignore these other influences is bad economics, and can only result in bad economic policies.

In fact, not all economists in the profession share this pluralist approach. The majority of them argue that there is no need to understand political science or sociology, or even to study history. According to this view, economics is about markets governed by a number of natural laws that are immutable

through time (and geography), and therefore independent of social conditions and institutions. In this view, economics is about studying individual behaviour and finding optimal solutions that are independent of history and time. Indeed, for most of the economics profession, the same economic theory should be able to explain problems in the United States (US) in 2020 as in Brazil in 1950. According to this view, the laws of economics are immutable and apply equally to all places and at all times.

The chapters in this book share a different view. They propose an alternative approach to economic theory – and therefore to economic policy – by taking into account socioeconomic conditions and institutions as they evolve through time: the problems in Brazil today are not the same as in the United States; the institutions in Europe are not the same as in North America. To think otherwise can only be dangerous, if we apply these policies to the real world. So, contrary to what Margaret Thatcher once said, that 'there is no alternative' (TINA) to the dominant doctrine, this book offers not only a credible alternative, but also one that provides a realistic view of how the real world actually works (Box 1.1).

While we discuss this more fully below, the mantra of TINA sends a dangerous message, and those who believe in its wisdom are doomed to repeat the mistakes of the past. A striking example of this flawed logic is the economic and financial crisis that began in 2007, which bears many similarities to the 1929 crisis that induced the Great Depression, hence why many have nicknamed it the Great Recession or the Lesser Depression. The similarities are striking not only in the statistics, such as the unemployment rate, income inequality or the rate of economic growth, but also in the response that many governments

BOX 1.1

TINA

TINA is a popular acronym that stands for 'there is no alternative'. It was the political slogan of British Prime Minister Margaret Thatcher, who in the early 1980s used it to justify economic policies of extreme austerity. Its meaning is clear: for Thatcher, market economies – or capitalism, (neo) liberalism, or laissez-faire – are governed by only one set of rules. Because of this, she had no choice but to impose draconian economic policies. The implication is that there is only one valid theory of economics, namely the neoclassical approach, and all other approaches are not acceptable. The purpose of this book is to show that there is an alternative to the dominant approach, valid in its scientific rigour and credible in its policy prescriptions.

BOX 1.2

THE GENERAL THEORY OF EMPLOYMENT, INTEREST AND MONEY

Published in 1936, *The General Theory of Employment, Interest and Money* remains the most influential book in economics of the twentieth century. Written at the height of the Great Depression, this book points out that the problem with contemporary economies is the lack of aggregate demand. In this situation, when markets are failing, the government has an important role to play in promoting economic growth, by spending (fiscal policy) and thereby contributing to the growth of aggregate demand. On 1 January 1935, Keynes wrote to his friend George Bernard Shaw that:

> I believe myself to be writing a book on economic theory which will largely revolutionize – not I suppose, at once but in the course of the next ten years – the way the world thinks about its economic problems. I can't expect you, or anyone else, to believe this at the present stage. But for myself I don't merely hope what I say – in my own mind, I'm quite sure. (Keynes, 1972a, pp. 492–3)

The General Theory of Employment, Interest and Money has a lasting influence on economic theory and policy today. See Chapter 4 for a portrait of Keynes.

provided at the beginning of the crisis. As in the Great Depression, when faced with dire evidence of the depth of the Great Recession in 2009 governments initially responded by increasing public spending, which had the immediate success of stopping the downward economic spiral. This was a triumph for good old Keynesian stimulus policies, prompting British Keynes biographer Robert Skidelsky to publish a book in 2009 entitled *The Return of the Master*, thereby highlighting the relevance of Keynes and Keynesian economics some seven decades after Keynes published his *magnum opus*, *The General Theory of Employment, Interest and Money* (Keynes, 1936) (Box 1.2).

However, the success of Keynesian fiscal stimulus created another problem: many policy-makers and economists interpreted its success as the end of the Great Recession, and proclaimed that economic growth was thus right around the corner. As a result, they quickly advised governments to begin cutting back on expenditures, otherwise their policy would prove inflationary or destabilizing. Government leaders agreed, and 2010 marked the reversal of public spending.

Indeed, once the economy appeared stabilized in 2010, several governments, under pressure to rein in their expenses, quickly abandoned their expansionary policies and started to cut expenses, and in some countries in very drastic ways. The fear was that high public deficits or debt would eventually cripple

BOX 1.3

EXPANSIONARY FISCAL CONTRACTION

The 'expansionary fiscal contraction' hypothesis assumes that the reduction of government spending and the adoption of austerity will have beneficial effects on economic activity, since these policies will affect both households' and firms' expectations. As such, agents will now form new expectations regarding fiscal expenditures and taxes in the future, and as a result will increase today their consumption and investment activities. These views are based on the ideas that increased government spending crowds out private sector spending. As such, reducing public spending will lead to increased private sector spending.

BOX 1.4

SECULAR STAGNATION

Secular stagnation is an expression that goes back to the writings of Alvin Hansen (see Box 1.5). In general, it refers to a long period of low economic growth, or rather a prolonged crisis of aggregate demand. The modern revival of the expression is credited mainly to Lawrence Summers, the former US Secretary of the Treasury. Today, many economies seem to be stuck in this secular stagnation, with growth rates hovering at best around 2 per cent yearly.

the economies, and threaten the fragile recovery. The logic now was that cutting fiscal expenditures would contribute to sustained economic growth and lead to prosperity.

This is what we call 'austerity policies': the idea that economies can grow from cutting government expenditures. Some have also called this approach 'expansionary fiscal contraction' (Box 1.3).

Yet, more than a decade after the official end of the crisis, we are still barely in the recovery phase of the expansion, and many parts of the world are threatening to slip back into recession. For instance, the European Union, and particularly the euro area, is gripped by a severe crisis of its own, and both the US and the Canadian economies are growing at very low rates, prompting many to ask whether we are in a situation of 'secular stagnation' (Boxes 1.4 and 1.5), which is further aggravated by the economic consequences of the COVID-19 pandemic.

BOX 1.5

ALVIN HARVEY HANSEN

Alvin Hansen (1887–1975) was an American economist, and professor at Harvard University, and was credited with introducing the views and ideas of John Maynard Keynes to the United States. In the late 1930s, he argued that 'secular stagnation' best described the American economy. Yet, the sustained economic growth that began in the early 1940s contributed to the concept being largely forgotten, until it was revived more recently.

In fact, history teaches us many things, and if economists and governments had studied the Great Depression and its aftermath, perhaps we would not still be underperforming, more than a decade after the 2007 crash, because the reaction governments had in the recent past is not dissimilar to how governments reacted in the aftermath of the 1929 financial crisis and the following Great Depression.

This analysis poses many challenges for economists and policy-makers. In this sense, economists must be able to better understand an increasingly complex world. This is what Keynes meant when he said that economists must be a bit of everything: sociologist, philosopher, mathematician, and so on. In fact, recently, Andrew Haldane, the chief economist at the Bank of England, mused that economists can learn from listening to different opinions and disciplines, to expand their 'wellspring of creativity'. Describing his own approach to economics, he says 'I have been begging, borrowing and stealing ideas over the last 15 years, to try and make sense of the world' (Clark, 2018).

This pluralist message is being heard more and more, and while it is a slow uphill battle, it is nonetheless the right path to follow for the common good. The economists contributing to this book have studied the real world long and hard, and arrived at the conclusion that we must rethink the old ways. US economist Alfred Eichner (1983, p. 238) once wrote that:

> This situation in which economists find themselves is therefore not unlike that of many natural scientists who, when faced with mounting evidence in support of first, the Copernican theory of the universe and then, later, the Darwinian theory of evolution, had to decide whether undermining the revelatory basis of Judeo-Christian ethics was not too great a price to pay for being able to reveal the truth.

While many economists seem to be in denial, and refusing to recognize the limitations of conventional thinking, a growing number of economists are recognizing the limits of the old ways – as are students. For instance, there is currently a growing movement among students around the world called 'Rethinking Economics', the purpose of which is precisely to introduce more pluralism within the teaching of economics. In France, Les Economistes Atterrés ('The Appalled Economists') is a group of faculty and students who are leading the fight to introduce more pluralism within economics. Other movements exist elsewhere as well. And at Harvard University, in November 2011, students walked out of Gregory Mankiw's economics class, over what they interpreted as a 'conservative bias' in his economics. Indeed, many claim that the economic theories and policies advocated by Professor Mankiw, among many others, were at the root of the 2007–08 crisis. The students at Harvard University were aware of this, and they left the class, demanding changes.

Keynes had warned us about such wrong policies. As he tells us in the single-paragraph first chapter of his most famous book, *The General Theory of Employment, Interest and Money*, 'the characteristics of the special case assumed by the [neo]classical theory happen not to be those of the economic society in which we actually live, with the result that its teaching is misleading and disastrous if we attempt to apply it to the facts of experience' (Keynes, 1936, p. 3). This book is largely based on that warning.

The role of ideology in economics

Understanding the world around us requires us to interpret what is going on, and that is not an easy task because it is always obfuscated by ideology. As we discuss below, as well as in the next chapter, there are two overall visions about how the economy works, or two ideologies, and whichever one you adhere to will taint the way you see the world around you (see also Chapter 3). These ideologies are in direct opposition to one another, and there are tremendous social forces and vested interests that seek to ensure the continued dominance of one over the other.

In essence, do you see markets being better off left to themselves and without any interference from the government? This is the 'laissez-faire' approach or the (neo)classical approach which Keynes referred to above: leave markets alone, minimize or even eliminate all government 'interference', which can only make things worse. Markets have some built-in stabilizers that ensure markets on their own are able to return to equilibrium.

We can illustrate this view using the analogy of a bowl and a marble. Once you drop the marble in the bowl, after moving around the marble will eventually gravitate toward the bottom of the bowl and then will come to a stop or a position of rest, that is, in equilibrium at the bottom. In this sense, this equilibrium can be defined as a position of gravitation: forces are pushing the marble to this position. What we must not do, therefore, is put anything in the path of the marble. On its own, without help, the marble will eventually return to its position of equilibrium.

If the economy suffers an 'exogenous shock' (kicking the bowl, so to speak), then the marble once again begins to move around, yet once the shock is over, the marble will eventually, once again, move back to the bottom of the bowl. To be clear, the marble never has any tendency to jump out of the bowl (this would be a crisis): the forces always dictate that the marble will always find its way back to equilibrium. What is important is to remove all obstacles in the path of the marble. 'Let it do its thing', and it will go back to equilibrium on its own.

As such, according to this approach, governments and government legislation often become obstacles that prevent markets from reaching their equilibrium. In other words, governments become the cause of economic recessions and depressions when imposing regulations, high taxes, tariffs, or by spending too much. Unions are also to blame, since they demand higher wages for their workers than what markets dictate and support. The end result is that resources are misallocated or misused and wasted, and that leads to unemployment, inflation, and slow economic growth or recessions.

From this philosophy are derived economic policies aimed at eliminating public deficits (balanced budgets legislation, for instance) as well as policies aimed at lowering taxes, and making it more difficult to join labour unions. The overall aim is to reduce the influence of institutions, including the state, in the workings of markets, which are seen as imperfections. As stated before, the vision here is of an economy in which forces of stability will move the economy naturally to its path of equilibrium. 'Stability' and 'convergence' are the two operative words.

The second approach is completely different – opposite in many respects – and sees markets, when left to themselves, as prone to instability, excesses and even crises: this is a dominant theme of this book. This contrasting approach, dominated by instability and chaos, argues that there are no inherent forces within markets that would enable economies to grow on their own for prolonged periods of time, or to push the economy back to an 'equilib-

rium' position. As such, markets need a little help to ensure the forces are such that they do not contribute to an economic crisis.

Using the marble and bowl analogy once more, imagine now that the bowl itself is constantly being kicked around, such that the marble is moving around wildly. It is no longer clear what are the long-run tendencies of the marble and, in fact, the marble may now jump out of the bowl. In this environment, chaos and instability are the operative words. And because of this, there is a role for the government to stabilize markets.

The difference between these two approaches largely comes down to how we see the role of the state: as a force for good or for evil? As stated above, if one sees markets as stable and converging naturally to an equilibrium, then governments are intrusive; naturally any involvement of the public sector will be seen as an assault on the wisdom of free markets. But others see governments in a positive light, as a way of helping markets to overcome some of their excesses and problematic behaviour. In this sense, the role of government is paramount. In other words, the role of the government is to make sure to put a lid on the bowl in order to prevent the marble from escaping.

In fact, without governments, contrary to mainstream thinking, markets are prone to periods of great instability: there are no stable forces leading the economy to an equilibrium. One of the reasons for this is that we live in a world of uncertainty, a central theme of John Maynard Keynes's thinking; a world in which we cannot predict the future. As he writes, 'we simply do not know' the future. Uncertainty has important consequences on how individuals and firms behave. Indeed, the 'extreme precariousness' of our knowledge of the future, but also of our understanding of how markets work, can have undesired consequences. Despite this, we somehow take decisions every day. Keynes called 'animal spirits' that little voice that somehow guides us in making these decisions in a framework of uncertainty.

In light of this climate of uncertainty, pessimism and optimism regarding the future are key in understanding what drives individuals and firms to act. For instance, when there is too much pessimism about the future, firms may not want to invest or seek funding from banks, which themselves may not want to lend. Consumers may not want to spend, but instead wish to accumulate savings. All these forms of behaviour contribute to depressing total demand on the market for produced goods and services, and lead to an inevitable crisis in aggregate demand.

This is where the government comes in. In times of great uncertainty, when aggregate demand is weak, governments can have a stabilizing influence on markets. When we are in a recession, and there is great uncertainty about the future of the economy, the government can enhance economic activity by means of an expansionary fiscal policy, that is, increase public spending. By purchasing goods and services from the private sector, or by transferring money to consumers, governments contribute to increasing total demand on the product market. In turn, this helps to create an atmosphere of confidence or optimism (or less pessimism), which then will allow firms to want to invest again and banks to lend (see Chapter 4 for a description of what is called the 'monetary circuit' and the role of banks). Hence, economic agents act according to 'sudden bursts of optimism and pessimism', as Keynes tells us.

In most universities the first approach is taught, and free market economics is presented as the only credible approach. Students are never asked to question this approach. They learn it, and then are quizzed on it, and eventually must accept it in order to graduate, at which time they go off into the private or public sector to work as economists, where they apply the lessons learned. If they become graduate students, they must write a thesis using this world view, or risk not receiving their degree. And once they become an economics professor, since they know nothing else, they also teach it. This is how the circle perpetuates itself. Yet, as Nobel laureate Joseph Stiglitz (2002) wrote, economics as taught 'in America's graduate schools . . . bears testimony to a triumph of ideology over science'.

This is a fundamental argument of this book. You will be introduced, in fact, to two overall approaches to economics, which we can label orthodox (or neoclassical or mainstream) and heterodox (or post-Keynesian). Each approach has its own assumptions and hypotheses, each is rooted in a given ideology, and each offers not only a very different interpretation of the real world but also vastly different theories and sets of policies to adopt in order to solve a number of economic problems. Each approach asks very different questions and hence provides very different answers.

Before we explore these two distinct approaches in detail, let us first discuss an argument that is of considerable importance: is economics a science?

Is economics a science?

In the sciences, knowledge is gained through observations and experimentation. For instance, we can test hypotheses in a laboratory, under specific conditions. We can recreate the conditions of space, or test the effects of the

lack of gravity on humans who are preparing to go in outer space – all within a laboratory. We can run computer simulations on how a new plane engine may perform, and test how changes in its design can affect performance.

Also, we know that there are immutable laws: the law of gravity, Newton's three laws of motion, the laws of thermodynamics, Einstein's law of relativity. And these laws do not change through time. For instance, gravity was the same 1000 years ago as it is today. Gravity may be different on Earth than it is on Jupiter, but the laws that govern gravity are specific. The same applies to the laws of motion: the planets, for instance, move in the same way today as they did billions of years ago. Light travels as fast today as it did in the past. Moreover, in the world of hard sciences, if the real world is complicated, we can always hold certain variables constant and isolate the effect of one variable on another. And if one experiment fails, we can repeat it as many times as needed until we obtain the desired results.

In economics, however, little or none of this can be done. For instance, we cannot recreate the conditions of the real world in a laboratory, such as the labour market or the banking system, and carry out tests to verify a hypothesis; nor can we simply hold certain real-world variables constant, or measure the specific effect of a single variable on the economy; what economists call the *ceteris paribus* condition, which translates into 'all other things being equal'. The real world is far too complicated; it is a place where everything is happening at once.

Of course, this does not mean that we cannot carry out tests or use models to advance our knowledge of how economic systems work; models can be very useful indeed. By definition, however, they are a simplification of the real world. As such, they must be used properly and with a series of explicit (rather than largely implicit) hypotheses. Further, these hypotheses must be rooted in reality and not simply *ad hoc* to satisfy the conclusions of the model. In other words, models must be representative of the real world. Given the discussion above (and see Chapter 2 for a discussion of current models used in economics), should models that assume a natural tendency towards equilibrium be used as a realistic representation of the real world?

In any economic system, societies and markets change all the time. The way cars are built today has nothing to do with how they were built in the days of Henry Ford; while government spending accounted for very little of overall gross domestic product (GDP) a century ago, today it accounts for far more. Today's institutions, like governments but also trade unions and corporations, do not resemble their former selves. Indeed, corporations

today are more complex and intricate than they were a century ago; a bank today performs very differently than in the past, not to mention the existence of shadow banks. As institutions change, they impact upon the way markets operate and perform, so a policy that may have been relevant at a time when agriculture accounted for an important part of the economy cannot surely be as relevant today where agriculture accounts for a relatively small part of overall economic output. This does not mean that there are no characteristics that persist: for instance, the way Keynes described the capitalist system in 1936 is still very much relevant today, and as such his 1936 book is still useful in understanding modern capitalism. But since the first task of economists is to observe the world around them, they must take note that the world today is characterized by the presence of large oligopolistic firms, large governments and the dominance of finance. This is not the economy of the early twentieth century.

These observations are important in deciding the use of models. Obviously, models must reflect how markets evolve and change over time. In using models, we must consider carefully the assumptions and hypotheses that are made: how much do they reflect the real world? And just as crucial, how important are the variables we exclude? Would their inclusion change the conclusions of the model and if so, how and why?

Regarding the changing nature of institutions, the question we need to ask, as hinted above, is whether institutions today are the same as in the past; and if not, can this be a sufficient reason to rethink economics? After all, does it really matter that institutions change? Could we not simply ignore these changes and go on analysing markets as if everything remained the same through time?

Change is therefore a central theme of our economies, and therefore there can be no universal laws in economics, like the law of gravity or the laws of motion in physics. As a result, economics is not a hard science. Given these changes, among many others, we cannot expect the same theories of economics from a century ago to still be relevant today. As a result, economic policies that may have been successful in the past may simply be wrong today.

This does not mean that economists cannot behave like scientists and observe the world around them. In this sense, economics is scientific, and is based, as Keynes tells us, on a 'vigorous observation of the real world'. The scientific method goes from observations to theorizing; going, however, much beyond surface phenomena. So there is a definite scientific approach within economics, as in other social sciences.

While this discussion may appear sensible to the vast majority of us, in reality it is shared today by only a very small percentage of economists and policy-makers. As stated above, at the time of writing, the vast majority of the profession believes in a very narrow definition of economics, one that we hope to show you in this book is wrong. Economists who believe in that approach believe in the laws of economic theory, that is, in the idea that economic theories can be applied anywhere and at any time. But what happens when the real world does not behave like the theories predict? What is wrong? In physics, when the real world does not live up to the theory, the theory is judged to be wrong and is eventually replaced. This is what happened, for instance, with the notion that the Earth is flat, or that the Sun rotated around the Earth.

As such, in economics, mainstream economists do not accept the scientific approach: despite the failure of many of their theories and policies, they still believe their theories are correct. They lay blame rather with institutions such as the state and unions for interfering with the laws of markets. The idea that their theories may be wrong is simply not a possibility in their minds. Two wonderful quotes by Keynes illustrate this picture well. First, in *The General Theory*, Keynes (1936, p. 16) compared this approach to 'Euclidean geometers in a non-Euclidean world who, discovering that in experience straight lines apparently parallel often meet, rebuke the lines for not keeping straight'. This is what mainstream economists do: they blame the real world for not behaving like their theory predicts. Second, if only they followed Keynes's wise words: 'When the facts change, I change my mind. What do you do, sir?'

The use of models and of mathematics

Let us discuss the role that models play in economics a bit more. Economists will often use tables and graphs to interpret and explain the world around them, or use sophisticated mathematics and statistical analysis. This is part of the economist's bag of devices. This does not negate the above discussion. In using models and mathematics, however, one must be careful to use them properly. Also, the absence of mathematics or sophisticated models does not render a theory useless. Economists must first be able to tell a story and explain the real world, and only afterwards should they rely on models and mathematics to support their conclusions, if the case warrants it.

Models are a simplification of reality, and as such they do not have to be complicated, but they must be realistic in the sense that they are meant to be a simplification of the real world, not a simplification of some fictitious world.

Models are rooted in assumptions, and it is these assumptions that must be a satisfactory reflection of the real world. If the assumptions are wrong, the model becomes useless. Imagine trying to explain the rotation of the Earth or the movement of planets based on the hypothesis that the Earth is flat, and that all celestial objects rotate around it.

For instance, our economies grow through increases in the demand for goods and services. As such, demand plays a central role. Any model that does not give demand such a role cannot be taken seriously. As another example, we live in a money-using economy where wages are paid in money, bank loans are made in terms of money, and money is what we use to buy goods. Hence, money must be at the very core of our economic models, and if this is not the case, the model must be cast aside in favour of one that integrates money in its analysis.

Some other common assumptions that we find in the most celebrated models are that full employment always prevails, unemployment is only temporary, money does not exist at all, economies evolved from barter, time is finite in the sense that models work only over a few periods, there is no accumulation of capital (that is, no investment), there exists only one good, and so on. Now, you tell us if these are realistic assumptions. Economist John Kay (2011) remarked quite appropriately that '[s]uch models are akin to Tolkien's Middle Earth, or a computer game like Grand Theft Auto'. Strangely enough, mainstream economists will admit that their assumptions are not realistic, but for them this is of no importance insofar as their models are internally consistent.

A perfect example of the dangers of using bad economic models is the total failure of mainstream economics in predicting the financial crisis that erupted in 2007–08. Indeed, not a single mainstream economist saw it coming. In fact, a few months before the crisis began in August 2007, the then Chairman of the US Federal Reserve, Ben Bernanke (2007), stated in his Congressional testimony in March 2007 that, '[a]t this juncture, the impact on the broader economy and financial markets of the problems in the subprime market seems likely to be contained'. Less than a year later, on 10 January 2008, he boldly claimed that '[t]he Federal Reserve is not currently forecasting a recession' (Bernanke, 2008).

The crisis was a proof of the complete collapse of mainstream economics. This prompted even the Queen of the United Kingdom, on 5 November 2008, to ask the question: 'Why did nobody notice it developing?' (Davidson, 2015, p. 1).

The answer is quite simple: nobody noticed the crisis because most economists are working with models that are simply wrong, and based on a series of flawed assumptions. In fact, in mainstream models, crises cannot occur, because markets are thought to be efficient. These models may acknowledge problems in one market or another, but not a systemic failure of the whole system at once. Yet, as 2007–08 has shown us, this can happen, and happens from time to time.

With respect to mathematics, like models, it plays an important role, but one that should be subservient to the story. In other words, economics is about telling a story of the world around us. This story must first be told, and once told, mathematics may be used to support that story.

The problem, however, is that today models and mathematics have taken over the economics, and economists rely too much on mathematics. As a result, the economic story gets lost. The increasingly more sophisticated models and statistical techniques become the focus of research, rather than economics. In a sweeping rebuke of the profession, economist Deirdre McCloskey (2005, p. 85), referring to the increasing mathematical nature of economics, wrote that '[i]f I am right in my criticism of economics – I pray that I am not – then much of what economists do nowadays is a waste of time'. This is precisely what we were pointing out above: the story of economics has been lost and replaced with mathematical sophistication. We need to go back to telling the story, which is what this book does.

Indeed, this is not only a recent story. In fact, Keynes, while recognizing the importance of mathematics in a supportive role, also warned us about the dangers of placing too much faith in mathematics. In *The General Theory* he wrote that:

> Too large a proportion of recent 'mathematical' economics are mere concoctions, as imprecise as the initial assumptions they rest on, which allow the author to lose sight of the complexities and interdependencies of the real world in a maze of pretentious and unhelpful symbols. (Keynes, 1936, p. 272)

Note that Keynes is precisely referring to the fact that many economists are more interested in making more sophisticated models than in explaining the real world. In terms of models, as Keynes (1973b, p. 296) wrote in 1938 in a letter to his friend and colleague Roy Harrod:

> Economics is a science of thinking in terms of models joined to the art of choosing models which are relevant to the contemporary world. It is compelled to be this,

> because, unlike the typical natural science, the material to which it is applied is, in too many respects, not homogeneous through time.

Joan Robinson (1962, p. 21), a colleague of Keynes at the University of Cambridge (see Chapter 15 for a portrait of her), explains that 'it is the business of economists, not to tell us what to do, but show why what we are doing anyway is in accord with proper principles'.

Economics is indeed the art of choosing models that are relevant to the world we live in, which is changing constantly. It must respect 'proper principles'. The task is not an easy one, of course, but that is the challenge economists set themselves, and that too many simply ignore.

Hence, better modelling (meaning more mathematical sophistication) is not the same as better economics. Today, there is a lot of energy being invested into making better and more sophisticated models, with an increasing degree of complexity, which is somehow supposed to help the economist in their task of evaluating the economy and predicting the future. Yet, the complexity of the model is irrelevant if its assumptions are flawed or wrong. Models must respect these 'proper principles'.

Economics and the social sciences

Where does economics fit then? It is, as Keynes (1973b, p. 296) tells us, 'essentially a moral science and not a natural science'. Before looking at how economics is actually related to other social sciences, we must first understand the world that economics is trying to analyse. In other words, what is this real world we live in? This should make clear why economics is right at home in the social sciences.

The approach or vision of economics in this book can be traced back a few hundred years to what is called the 'classical' period. This is the time of such great economists as Adam Smith, David Ricardo and Karl Marx; a time when agriculture was a dominant component of economic activity. While it is often the tradition to see vast differences between these three economists – and undoubtedly there were important differences – they all shared a similar approach to economics: they all saw mid-eighteenth-century society in terms of a struggle between social classes. Indeed, society was divided into the capitalists, that is, those who owned the means of production and in particular owned the tools to till the land and grow crops; the rentiers, or those who rented the land to the capitalists; and the workers, that is, those who worked in the fields.

This way of seeing society as comprising social classes or macro-groups leads inevitably to the possibility of conflict: the potential conflict between classes as to the division of wealth. If you grow a sufficient amount of corn to guarantee crops the following year, what do you do with what is left over, or what is called the surplus? In other words, what do you do with the surplus corn over and above what is needed for planting the crops next season? Who gets what? How is it divided between capitalists, rentiers and workers?

After all, rentiers own the land on which the crops grew. So are they not entitled to the surplus? Without their land, there would not have been any crops to begin with. Yet, capitalists are the ones who own the machines needed to ensure the crops are planted and harvested, so surely they should get the surplus? But what about workers? After all, they are the ones who do the actual, physical work. Without workers to do the hard work, there would be no corn. As you can see, everyone claims a part of the surplus, and in this distribution of the surplus lies a potential conflict.

This way of seeing society is still relevant today. We still have capitalists who own factories and companies; the rentiers, or what we often call financial capitalists; and of course, we still have a large class of workers. Inevitably, there is still the great potential for conflict. Workers and their unions are again under attack, and increasingly so. Corporations are making record profits, yet workers' wages are stagnating if not declining in real terms.

Related to this whole discussion is the debate over the increased polarization of wealth: the growing discrepancy between the very rich and the poor. This is the great conflict over the distribution of income and wealth. This has been made strikingly clear in recent years with social movements such as Occupy Wall Street, which drew attention over what has become known as the '1 per cent'. At the core of this anger is the fact that the wealth is becoming increasingly concentrated at the top, among the 1 per cent and even among the 0.1 per cent of the population, while the rest struggles to make ends meet.

As this book will show, the distribution of income is an important component of overall economic growth. If the distribution of income or wealth is skewed, then this will affect economic growth in a negative way. There are many questions regarding this problem. What causes inequality? How is it precisely related to economic activity and growth? How do we solve the problem? Indeed, there is perhaps no more pressing matter today than the question over the inequality of income and wealth distribution within, as well as between, countries.

These are not easy questions to answer. But one thing is clear: to answer them, we must understand how economic policy is made, and how social dynamics come into play. This requires understanding what goes on beyond the confines of a narrowly defined notion of economics. As Keynes argued, we need to understand a bit of everything. So how then is economics related to other social sciences, and what can they teach the economist?

Political science can teach us many things about how economic policies are actually formulated and adopted. Economists can only recommend policies: politicians are the ones who implement them, so an understanding of how political parties and governments operate, and how policies are ultimately adopted, is paramount. Why is one policy adopted rather than another? Sometimes it has less to do with what is right for the country than what is the right thing to do for a political party to get re-elected. This will often have to do with the ideological leanings of the political party in power, and whether its representatives believe in free markets. As far as political science can help us to understand power relationships between individuals and groups, the economist must be able to understand these relationships when they provide policy advice.

Psychology is another of the social sciences that has much to do with economics. For instance, economists are very interested in why and how consumers spend their income, and why firms decide to invest. Keynes famously referred to investment decisions by firms as being influenced by 'animal spirits'. According to Keynes (1936, pp. 161–2):

> Even apart from the instability due to speculation, there is the instability due to the characteristic of human nature that a large proportion of our positive activities depend on spontaneous optimism rather than mathematical expectations, whether moral or hedonistic or economic. Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as the result of animal spirits – a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities.

Hence, understanding 'human nature', as Keynes puts it, is an important aspect of the economist's work. But what governs these animal spirits and these bursts of optimism? If not based on mathematical calculations, we must understand the motivations of the mind, which may require us to know the psychology behind such motivations.

Also, psychology helps us to understand motivations, the difference between needs and wants, full rationality versus bounded rationality, the impact of

cognitive dissonance on decision-making, and more. It is a growing area of interest for economists looking to understand how decisions are made in an uncertain and complex world.

There is much to learn from sociology as well, especially with respect to the behaviour of social classes and movements, the exercise of power, the role of institutions, and the evolution of capitalism. In his famous book *The Great Transformation*, Karl Polanyi (1944) discusses the concept of 'embeddedness', that is, the notion that individuals and firms are part of a greater existing social structure, and the actions between them must be analysed within these social networks.

Finally, history has probably the most important influence on economists. For instance, the economic and financial crisis that began in 2007–08 shares many characteristics with the Great Depression of the 1930s. Indeed, when the crisis erupted in 1929, income inequality – that is, the difference between the rich and the poor – was strikingly similar to the inequality that existed in 2007, thereby suggesting that in both cases inequality had an important role to play. Moreover, like in 1929, in the aftermath of the current crisis, governments adopted expansionary fiscal policies in the hope of turning fortunes around and leading us to a path of recovery. Yet, like in the Great Depression, when the governments stopped their fiscal largesse the economies slowed down and threatened to bring about another round of recession. Studying history should lead policy-makers to avoid repeating the mistakes of the past; mistakes that are unfortunately repeated today, owing to the short-sightedness of most of the economics profession.

What then is economics?

In the previous sections we explored how economics is influenced by ideology, and what it shares with other social sciences. But we have not yet discussed what economics is. What then is economics?

This is not an easy question to answer. It depends on who you ask. As has become clear by now, there are two very different approaches or visions in economic analysis, and these two very broad schools of thought adopt very different methodologies; they ask very different questions and provide very different answers.

We can call these two approaches by two very general names – 'orthodox' and 'heterodox' – although there may also be other names. For instance, orthodox economics is often referred to as neoclassical economics or

mainstream economics. Today, the large majority of the economics profession shares this approach. Heterodox economics is an approach that not only rejects orthodox economics and its presuppositions, but also proposes a clear alternative. It can also be labelled post-Keynesian or even post-classical economics.

If you ask an orthodox economist to define economics, you will get a very different answer than if you ask a heterodox economist. This should not be a surprise to you by now. After all, as we have tried to make clear, different visions entail a different definition of economics. Chapter 3 will develop the peculiarities of each school of thought, but let us focus here on a definition of what economics is, or rather: what does the economist do?

According to orthodox economists, the usual answer is that economics is about the 'efficient allocation of scarce resources'. There are three important words here: 'efficient', 'allocation' and 'scarce'. According to this approach, scarcity plays a central role: when economists say resources are scarce, they mean that their supply is limited. All resources or goods are scarce, irrespective of whether you are discussing labour, capital goods, water or money.

Since the supply is scarce, it must therefore be allocated carefully among the various demands for the good. Hence, the supply must be rationed. We cannot allocate too much of one resource to a particular market, because then there will be too little in other markets.

This brings us finally to the 'efficient' component of the above definition. It is assumed that markets are the best way to allocate the scarce resources among the competing demands for them. The laws of the market, which are deemed supreme, are entrusted with this allocation. It is assumed that markets do not harm; on their own they will achieve this efficient allocation. In this sense, there should be no interference with the laws of markets, and any institution, say the state or trade unions, but also price boards and more, that intervene with markets must be eliminated or their influence seriously curtailed.

Heterodox economists, however, consider economics in a very different way. First of all, rationing existing supplies of resources implies that markets are always operating at full employment. Yet, this is not the case, as economies typically do not perform at full capacity. In the real world, there is always some slack such that an efficient allocation is not required (see Chapter 9). For instance, the labour market usually has a 'reserve army' of unemployed. There is therefore no need to allocate labour efficiently between one sector

and another. Firms typically perform at less than full capacity, often around 85 to 90 per cent, such that there is always room in the short run to increase production, at least a little bit, to meet increased demand on the market for produced goods and services.

Economists must therefore try to explain why markets do not function at full capacity or full employment, and propose policies that will resolve these problems. Neoclassical economists usually attribute these problems to 'market failures', as if correcting for these failures will automatically bring economies to full employment. Such market failures are usually attributed to the intervention and interference of governments and other institutions.

In contrast, heterodox economists see these problems as inherent in the way markets operate: markets are complex, and how they operate must be properly and adequately explained. Heterodox economists must therefore explain how production occurs, how labour is integrated into the process of production and accumulation, how wages are determined, and how profits are generated. What are the requirements for economic growth?

In this sense, economics is the study of the dynamic process of production, accumulation and distribution within the context of existing social and institutional relations, and of the requirements for economic growth, while keeping in mind that these processes are subject to periods of instability, the forces of which must be explained.

Micro- versus macroeconomics

Before we end this introductory chapter, let us discuss briefly the differences between microeconomics and macroeconomics. Simply put, microeconomics (from the Greek meaning 'that which is small') is about the behaviour of individual agents, households, firms, and markets or industries. Government policies are also analysed for their impact on individual agents.

Macroeconomics, on the other hand, is about the economy as a whole, and how economic policies impact upon the overall economy; more specifically, growth, unemployment, distribution and inflation. Neoclassical or mainstream economics believes that macroeconomics is simply an aggregation of individual behaviour. In other words, the roots of macroeconomics are to be found in microeconomics, that is, what economists call microfoundations (see the Conclusion of this volume for a full discussion). In this sense, there is no need for macroeconomics. After all, if we can understand the economy as a whole based on the study of a 'representative agent', then why bother

with macroeconomics? Hence, by studying the behaviour of one agent we can extrapolate and understand the whole economy.

Yet, macroeconomics is subject to laws of its own that are unrelated to microeconomics. There are a number of fallacies of composition, or paradoxes, according to which what may be good for an individual or a firm may be harmful for the economy as a whole. A few examples will shed some light on this.

Take, for instance, savings. It makes sense for individuals to save a fraction of their income for a rainy day. After all, increased savings gives us security in case something unexpected happens. Yet, when we save, we are obviously not consuming. So, assume that everyone saves, then firms would be unable to sell many of their goods or services, and the economy would suffer. Increased savings for all may result in unemployment. This is the paradox of thrift.

Macroeconomics is more than just the study of the economy as a whole, as it also analyses the laws of production and distribution that govern it. By focusing on social groups rather than individuals, macroeconomics is about the dynamics of power relationships and the hierarchy of these groups relative to others, and the place of institutions. These questions cannot be treated by microeconomics, such that a field of its own, namely macroeconomics, is required to do justice to the complexity of these relations.

Finally, we live in a money-using economy, so, as Schumpeter argued, money should be introduced at the very beginning of economics (Chapter 4 discusses this argument in more detail). That is why this book, in contrast to all other macroeconomics textbooks, begins with an explanation of money (Chapter 4), the banking system and finance (Chapters 5, 6 and 7), rather than considering them later on, which is typical of mainstream textbooks.

REFERENCES

Bernanke, B.S. (2007), 'The economic outlook', Testimony before the Joint Economic Committee of the US Congress, Board of Governors of the Federal Reserve System, 28 March, accessed 22 October 2015 at www.federalreserve.gov/newsevents/testimony/bernanke20070328a.htm.

Bernanke, B.S. (2008), 'The economic outlook', Testimony before the Committee on the Budget of the US House of Representatives, Board of Governors of the Federal Reserve System, 17 January, accessed 22 October 2015 at www.federalreserve.gov/newsevents/testimony/bernanke20080117a.htm#fn1.

Clark, T. (2018), 'What Grayson Perry can teach the Bank of England', *Prospect Magazine*, 18 April, accessed 3 July 2020 at https://www.prospectmagazine.co.uk/magazine/new-thinking-at-the-old-lady-the-bank-of-englands-unlikely-new-advisers.

Davidson, P. (2015), *Post Keynesian Theory and Policy: A Realistic Analysis of the Market Oriented Capitalist Economy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Eichner, A. (1983), *Why Economics is Not Yet a Science*, Armonk, NY: M.E. Sharpe.

Kay, J. (2011), 'The map is not the territory: an essay on the state of economics', *John Kay: Accessible and Relevant Economics*, 4 October, accessed 24 August 2015 at www.johnkay.com/2011/10/04/the-map-is-not-the-territory-an-essay-on-the-state-of-economics.

Keynes, J.M. (1924), 'Alfred Marshall, 1842–1924', *Economic Journal*, **24** (135), 311–72.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Keynes, J.M. (1973a), *The Collected Writings of John Maynard Keynes. Volume XIII: The General Theory and After: Part I*, edited by D. Moggridge, London: Macmillan and Cambridge University Press.

Keynes, J.M. (1973b), *The Collected Writings of John Maynard Keynes. Volume XIV: The General Theory and After, Defence and Development*, London: Macmillan and Cambridge University Press.

McCloskey, D. (2005), 'The trouble with mathematics and statistics in economics', *History of Economic Ideas*, **13** (3), 85–102.

Polanyi, K. (1944), *The Great Transformation: The Political and Economic Origins of Our Time*, New York: Farrar & Rinehart.

Robinson, J. (1962), *Economic Philosophy*, Harmondsworth: Penguin Books.

Skidelsky, R. (2009), *The Return of the Master*, New York: Public Affairs.

Smith, A. (1759), *The Theory of Moral Sentiments*, London: A. Millar.

Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, 2 vols, London: W. Strahan & T. Cadell.

Stiglitz, J.E. (2002), 'There is no invisible hand', *Guardian*, 20 December, accessed 24 August 2015 at www.theguardian.com/education/2002/dec/20/highereducation.uk1.

A PORTRAIT OF ADAM SMITH (1723–90)

Adam Smith, born in Kirkcaldy, Scotland, is considered, along with David Ricardo and Karl Marx, as one of the fathers of classical political economy, a period that is also known as the birth of economics.

Within his vast contribution to economic analysis, Smith is known in particular for two important books, *The Theory of Moral Sentiments* (1759) and *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776), more commonly referred to simply as *The Wealth of Nations*. Smith was a leading member of the Scottish Enlightenment movement.

Smith began his studies in social philosophy at the age of 14 at the University of Glasgow, and in 1740 entered the University of Oxford. Through a series of public lectures at the University of Edinburgh in 1748, Smith befriended in 1750 colleague and philosopher David Hume, himself an accomplished writer. This friendship resulted in Smith returning to the University of Glasgow to accept a professorship in Logic in 1751, and then in Moral Philosophy in 1752. He would eventually become Rector of that university in 1787, although it was more an honorary position.

Smith wrote *The Theory of Moral Sentiments* in 1759, the premise of which is based on the notion that morality was influenced by the sympathy between individuals in society. The book was very successful and gave Smith great notoriety. He received a Doctorate in Law in 1762 from the University of Glasgow.

From 1764 to 1766, Smith left academia to tutor the future Duke of Buccleuch. This enabled him to travel widely through Europe, especially in France, where he befriended a number of intellectuals, notably François Quesnay, a leader of the Physiocrats. He retired in 1766 to his native Kirkcaldy, where he would eventually write *The Wealth of Nations*, a book that would take him more than nine years to write.

To this day *The Wealth of Nations* remains one of the most important books written in economics. The book is subject to different interpretations. Neoclassical economists tend to place the notion of the invisible hand at the heart of the book, emphasizing the virtues of free markets. For heterodox economists, by contrast, the book is more about the creation of wealth. Following the views of Quesnay and the Physiocrats, to generate wealth in the next period, the economy must produce an output that is capable of reproducing itself in the current period. Any surplus production must be divided among the various social classes, and this can generate conflict.

EXAM QUESTIONS

True or false questions

1. Economic models are a simplifications of the world in which we live.
2. Secular stagnation can best be described as a short-term disturbance in the ways markets operate.
3. There are no credible alternatives to the dominant approach in economics.
4. Ideology plays a significant role in the way economists see the world around them.
5. In economics, there are natural laws akin to the law of gravity in physics.

6. Fallacies of composition suggest that principles that apply to the microeconomic level also apply to the macroeconomic level.
7. Economics is the study of the dynamic process of production, accumulation and distribution within the context of existing social and institutional relations, and of the requirements for economic growth, while keeping in mind that these processes are subject to periods of instability, the forces of which must be explained.
8. Austerity is a policy by which governments reduce drastically their expenditures.
9. Pluralism is an approach to economics that emphasizes the importance of other disciplines in the understanding of the real world.
10. Laissez-faire economics is best described as a stable system in which forces will bring the economy to an equilibrium.

Multiple choice questions

1. To better understand the real world, economists:
   a) should be concerned strictly with microeconomics;
   b) should develop increasingly sophisticated models;
   c) should adopt a more pluralist approach;
   d) none of the above.
2. According to TINA:
   a) there is a multiplicity of valid approaches in economics;
   b) there is no alternative to mainstream economics;
   c) there are great benefits from adopting a pluralist approach;
   d) governments should always increase spending to spur economic growth.
3. What is the role of ideology in economics?
   a) Ideology plays no role in economics.
   b) It determines whether households should get higher wages.
   c) It determines the degree to which mathematics should be used in economics.
   d) It influences how we see markets operating and the appropriate government response.
4. The two best words to describe the mainstream or neoclassical view are:
   a) instability and chaos;
   b) stability and chaos;
   c) convergence and growth;
   d) stability and convergence.
5. The two best words to describe the heterodox or post-Keynesian approach are:
   a) instability and chaos;
   b) stability and chaos;
   c) convergence and growth;
   d) stability and convergence.
6. Governments:
   a) can play an important role in stabilizing markets;
   b) can stabilize markets only through austerity policies;
   c) are never able to stabilize markets;
   d) none of the above.
7. Uncertainty is understood as a situation when:
   a) the past is irrelevant to decision-making;
   b) all decisions are made with respect to current outcomes;
   c) present decisions can only be taken in light of perfect certainty;
   d) future outcomes are unknown: 'we simply do not know'.

8. Bursts of optimism and pessimism:
a) play no role in decision-making;
b) play an important role in determining investment decisions;
c) can safely be ignored regarding all aspects of decision-making;
d) are important only in times of recessions.

9. Models:
a) are to be used only occasionally;
b) should be based on unrealistic assumptions of the real world;
c) should be a realistic simplification of the real world;
d) have no place in economics.

10. Economics is about:
a) understanding the scarcity of all goods;
b) understanding the microfoundations of macroeconomics;
c) understanding how bank credit and money are related in a production economy;
d) understanding the immutable laws of markets.

2

The state of macroeconomics

Louis-Philippe Rochon and Sergio Rossi

OVERVIEW

This chapter:

- explains the current state of mainstream macroeconomics;
- argues that the dominant model for economic analysis must be abandoned;
- points out that the post-Keynesian alternative is more realistic.

KEYWORDS

- **Austerity:** An economic policy that aims at reducing public expenditures to support economic growth.
- **Financial fragility:** A situation where the whole economic system is exposed to a financial crisis owing to the volume of private debt with respect to disposable income.
- **Frictions:** Hindrances to the smooth working of the alleged 'law' of supply and demand in a free market economy.
- **Great Moderation:** The historical period (1985–2007) during which the global economy was characterized by low volatility in economic growth, and low rates of inflation and unemployment.
- **Microfoundations:** Analytical tools provided by microeconomics that are meant to be useful to understand the working of the economic system as a whole.

Why are these topics important?

Twenty-first-century mainstream macroeconomics is in disarray. This much we know since at least the global financial crisis that erupted in 2007–08, when mainstream economists stopped agreeing with each other about economic theory and, more importantly, policy, resulting in a civil war among the mainstream. This is in stark contrast to the situation that prevailed before

the crisis, when a large consensus existed. In fact, the crisis revealed the fundamental weakness of mainstream macroeconomics, and the degree to which the macroeconomic emperor has no clothes: economic theory fell apart and revealed the vast distance between it and the real world, and the inability of mainstream macroeconomics to propose relevant policies.

The new coronavirus (COVID-19) pandemic and the ensuing economic crisis of 2020 only exacerbated the divisions, further revealing the hypocrisy of many mainstream economists who have made it clear to everyone that austerity remained the best economic policy. Yet everybody has witnessed the consequences of years of austerity: our institutions today have been rendered too weak to respond adequately to the many problems we were confronted with during the COVID-19 crisis. The challenges we now face in rebuilding our institutions and economies are great, and the current state of mainstream macroeconomics is not up to the task of providing any valuable theoretical or policy insights, putting the future of economic prosperity in doubt. In fact, we need to discard the entirety of mainstream macroeconomics as irrelevant. After years of austerity, after decades of being told 'the state does not have the money', we now see through the lies perpetuated by the mainstream: around the world, we see that 'money' was always there; to wit, governments are not bound by a budget constraint. Governments refused to spend simply because they wanted to avoid entering into debt, thereby leading private sector agents, notably consumers, to increase their own debt. This induced the outbreak of the global financial crisis in 2007–08.

This crisis also revealed something very profound about mainstream economics. Not only were current models not able to predict that crisis, but economists were unable to propose appropriate policies to deal with (then) the largest crisis since the Great Depression induced by the 1929 financial crisis. Indeed, apart from a brief flirtation with fiscal expansion in 2009, starting in 2010 the economics profession proposed more austerity and lower interest rates despite the limited (if any) success of either policy. There existed a theoretical black hole, and mainstream economists, nurtured on their precious models, were simply intellectually unable to carry out a deep rethinking of their approach, which is now necessary urgently.

The same intellectual void applies to the COVID-19 crisis. While governments around the world have created very large fiscal deficits to deal with three simultaneous crises (the health crisis, the supply chain crisis and the crisis of aggregate demand), austerians, unable to free themselves of their intellectual straightjacket, are falling back on the only refrain they know: economies will collapse under the weight of public debt, and only massive

cutbacks will put economies back on the path to prosperity. 'Business-as-usual' economics is making a comeback.

The end result is confusion. How can anyone propose 'same-old, same-old' policies, when the world is revealing itself as being so different? As stated above, mainstream economists seem unable to go beyond their models and propose bold new ways of thinking about economics and economic policies.

While Keynes referred to a colossal muddle, an observation equally applicable today, Vines and Wills (2018, p. 2) admit that it is 'no longer clear what macroeconomic theory should look like'. Recently, Paul Krugman himself admitted that 'we can't just use standard macro models off the shelf' (Bloomberg, 2020).

These issues are important, because we need models that allow us to make relevant policy proposals. In a post-COVID-19 world, this is of great importance. How will our economies recover from these crises? What role can fiscal policy play? The potential loss in wealth for the so-called 99 per cent of the population as a result of the COVID-19 crisis is enormous. More than ever, we need relevant theories and even more relevant policies.

Before the crises

Today's civil war among mainstream economists is a far cry from the consensus that existed only a few years before the 2007–08 crisis. Indeed, at the 2003 presidential address of the American Economic Association, Nobel laureate Robert Lucas (2003, p. 1) declared rather triumphantly that the 'central problem of depression-prevention has been solved'. Given the curiously long period of economic growth the United States was experiencing at the time, the so-called Great Moderation – a period of some 20 years (1985–2007) characterized with low inflation rates and reduced volatility in real gross domestic product (GDP) growth – many economists thought advances in macroeconomic theory had contributed greatly to improved economic performance (Box 2.1). As a result, recessions were considered a thing of the past.

Lucas's (2003) statement echoed some remarks made a few years earlier, which appeared in a 1997 *Foreign Affairs* article entitled 'The end of the business cycle?', in which the author argued that:

> modern economies operate differently than nineteenth-century and early twentieth-century industrial economies. Changes in technology, ideology, employment, and finance, along with the globalization of production and

BOX 2.1

THE GREAT MODERATION

In economics, the Great Moderation refers to a period of a little over 20 years, during which the United States (US) economy, as well as others, experienced low inflation and reduced volatility in real GDP growth. In turn, this is said to have had great benefits: more stable employment, lower uncertainty regarding the economy, and better market coordination. In other words, because inflation is low and output is not fluctuating too wildly, economies flourish. As a result, economists thought (wrongly) that they had finally discovered the secrets to a well-functioning economy. They argued that better theories and policies were responsible for the Great Moderation. Of course, the 2007–08 crisis quickly put an end to that hypothesis.

> consumption, have reduced the volatility of economic activity in the industrialized world. For both empirical and theoretical reasons, in advanced industrial economies the waves of the business cycle may be becoming more like ripples. (Weber, 1997, p. 65)

Nobel laureate Paul Krugman (2009a) labelled this period 'the golden era of the [economics] profession'.

However, the Great Moderation proved to be not the end of recessions, but rather the calm before the economic storm: the Great Moderation gave way to the Great Recession, and depression economics was thus back as a dominant theme among industrialized countries, after having been a characteristic of developing economies.

The crisis, however, brought more than a simple reminder that economic systems are still subject to prolonged periods of low growth. It notably brought to the fore a great divide among economists as to the causes of growth and the role of policy. In other words, the Great Consensus that existed before the crisis evaporated. In its place, as Lavoie (2018, p. 15) states, 'there is considerable dissatisfaction with the current state of mainstream macroeconomics'.

To be sure, disagreements are par for the course and always existed among economists: Keynesians versus monetarists, or freshwater versus saltwater economists (Box 2.2).

Indeed, as Blanchard (2009, p. 213) claims, relations 'were tense, and often unpleasant'. Today, these new divisions seem to run much deeper, and go

BOX 2.2

SALTWATER VERSUS FRESHWATER ECONOMICS

Saltwater economists are associated with US universities whose campuses are on or close to the ocean coasts, with their salt sea waters. Among these universities are Harvard University, MIT, Yale University, Colombia University and Princeton University. These economists were generally thought of as Keynesians, such as Paul Samuelson, Robert Solow and James Tobin. In contrast, freshwater economics was practised by those at inland universities, usually close to the Great Lakes, such as the University of Chicago, Cornell University, Northwestern University and Carnegie Mellon University. Among the more notable economists were Robert Barro, Robert Lucas and Thomas Sargent. The division between them was noticeable. As Robert Hall wrote in 1976:

> As a gross oversimplification, current thought can be divided into two schools. The fresh water view holds that fluctuations are largely attributable to supply shifts and that the government is essentially incapable of affecting the level of economic activity. The salt water view holds shifts in demand responsible for fluctuations and thinks government policies (at least monetary policy) is [*sic*] capable of affecting demand.

But these differences largely disappeared with the years leading up to the Great Moderation. As Mankiw wrote in 2006 (p. 38), right before the eruption of the financial crisis:

> In macroeconomics, as the older generation of protagonists has retired or neared retirement, it has been replaced by a younger generation of macroeconomists who have adopted a culture of greater civility. At the same time, a new consensus has emerged about the best way to understand economic fluctuations . . . Like the neoclassical–Keynesian synthesis of an earlier generation, the new synthesis attempts to merge the strengths of the competing approaches that preceded it.

As we argue in this chapter, this consensus dissipated and gave way to less civility, or a 'civil war'.

beyond unpleasantness. In fact, since the crisis that erupted in 2007–08, something close to a civil war has been brewing among mainstream macroeconomists, as Krugman (2013) pointed out, with a cohort of well-known dissident economists now openly questioning the established doctrines and the use of the so-called dynamic stochastic general equilibrium (DSGE) model (Box 2.3). *The Economist* (2009) referred to this scenario as the 'turmoil among macroeconomists'.

While the chorus of dissent is still growing regarding the use of DSGE models and their accompanying microfoundations, as exemplified by the Spring–Summer 2018 issue of the *Oxford Review of Economic Policy* (which dedicated this double issue to a discussion of the inherent flaws of DSGE models), one must wonder how serious the economics profession is, as a

BOX 2.3

THE DSGE MODEL

The dynamic stochastic general equilibrium model is the dominant means of analysis in economics. It seeks to explain economic fluctuations based on a microeconomic, general equilibrium approach. The model assumes that the economic system can be modelled by simply aggregating the behaviour of individuals. It is firmly grounded in the concept of microfoundations to macroeconomics. It makes a number of assumptions about the economy, such as perfect competition, price flexibility and rational expectations, to which a number of frictions are added to generate disequilibrium outcomes. New Keynesians use a modified model in which prices are set by monopolistic firms, a variant of the neoclassical model. In the last decade, there has been growing dissent over these models for their unrealistic assumptions. They are at the heart of the discussion over why economists were not able to predict the 2007–08 financial crisis.

whole, about the future of such models and of macroeconomic theory in general.

The dissent over DSGE models is growing, but there is nevertheless considerable resistance. In this chapter, we examine closely this dissent – or the so-called 'civil war' discussed earlier – and whether it carries any consequences for the future development of macroeconomics, or whether it is just a microfoundations storm in a teacup. In doing so, we will argue that while the dissident critique is certainly welcome from a post-Keynesian perspective, it does not amount to a rejection of current thinking or models, but rather some tweaking around the mainstream edges. As King (2012, p. 150) explains, it becomes a debate over 'our [microfoundations] are better than yours'. Indeed, despite Vines and Wills's (2018, p. 15) admission cited above, there is still considerable agreement over what 'macroeconomic theory should look like'.

In other words, the DSGE model itself is not seriously questioned by the economics profession, and still seems to be the only legitimate model for it. The debate is over the appropriate microfoundations, and the push to include some imperfections, such as financial frictions. This is the position, for instance, of Olivier Blanchard (2014, p. 31), who argues that 'DSGE models should be expanded to better recognize the role of the financial system'[1] despite the fact that such models are 'loaded with questionable assumptions' (Blanchard, 2009, p. 225) and are 'seriously flawed' (Blanchard, 2018, p. 44). Nevertheless, the author believes that these models 'are eminently improvable and central to the future of macroeconomics' (Blanchard, 2018, p. 44).

BOX 2.4

MODELS AND POLICY

Recall the discussion in Chapter 1 about models and policy: models are what policy-makers use to develop and implement policies. Not only were economists not able to predict the 2007–08 financial crisis, but they were also unable to suggest appropriate policies. In this sense, the DSGE model was responsible for the great policy void. The fear is that by retaining them, economists will be unable to further assess the state of the economy, unable to predict the next great crisis, and unable once again to propose policies and contribute to economic prosperity.

Others disagree. For instance, Robert Solow (2008, p. 244) claimed that 'adding some realistic frictions does not make it any more plausible that an observed economy is acting out the desires of a single, consistent, forward-looking intelligence'.

The consequences of this debate are important for the future of economics and of economies. Recall that these are the models that were unable to predict the biggest crisis since the Great Depression. Keeping them, even by adding additional frictions to make them appear more realistic (they still are not), will not begin to address the problems of their unrealistic nature (Box 2.4). Are economists setting themselves up to repeat the past?

The inability or unwillingness of the economics profession to question the legitimacy of the DSGE model comes down to the absence, in the eyes of mainstream economists, of a better solution: 'there is still no alternative' (Box 2.5).

In this sense, the argument has come full circle, and in the end the recent brouhaha over mainstream models will amount to nothing much in terms of

BOX 2.5

TINA

An expression made popular by former British Prime Minister Margaret Thatcher, the saying that 'there is no alternative' is better known by its acronym, TINA. First used by nineteenth-century British philosopher Herbert Spencer, Thatcher used it in the 1980s to defend her market-friendly, yet draconian policies of austerity, claiming that there was no alternative economic theory or set of policies at her disposal. In other words, she had no choice but to adopt austerity policies.

reforms. The economics profession at large will still keep the DSGE model, although its adherents may introduce some minor amendments, and add some financial fragility to give it the semblance of realism. If successful, the result may prove to be the re-emergence of yet another synthesis where the core neoclassical ideas remain intact, but where crises are made possible through some sort of failure of the model. Once remedied, the economic system – and the economics profession – goes back to its business as usual.

This said, the only positive light might come from outside academia, from institutions that are looking beyond the DSGE model. We will briefly touch on this in the conclusion. Let us first present the traditional mainstream view of macroeconomic analysis.

The traditional mainstream view

The degree of consensus in macroeconomics that existed before the financial crisis was impressive, not only over the need for microfoundations and the use of models such as DSGE, but also in its rejection of Keynesian fiscal stimulus. In this sense, it is worth expanding upon Lucas's (2003, p. 1) earlier quote:

> Macroeconomics in this original sense has succeeded. Its central problem of depression-prevention has been solved, for all practical purposes, and has in fact been solved for many decades . . . The potential for welfare gains from better long-run, supply-side policies exceeds by far the potential from further improvements in short-run [Keynesian] demand management.

Echoing this Great Consensus or ideological *détente*, Blanchard (2009, pp. 211, 225), who at the time was Chief Economist at the International Monetary Fund (IMF), in an often-quoted paper entitled 'The state of macro' argued that there had been a 'broad convergence of vision' among macroeconomists: 'The state of macro is good . . . Macroeconomics is going through a period of great progress and excitement.' Similarly, a few years earlier, Chari and Kehoe (2006, p. 3), writing in the *Journal of Economic Literature*, stated that '[o]ver the last three decades, macroeconomic theory and the practice of macroeconomics by economists have changed – for the better', naming the use of DSGE models as a contributing factor.

This so-called 'progress' was evidenced, it is argued, by the Great Moderation, which in turn was attributed to good economic policies and to good economic models (theories), meaning theories that produced policies that stimulated the economy in ways that best responded to exogenous shocks (see

Davis and Kahn, 2008; Coric, 2011 for a survey). In other words, good policies reduced output volatility. Others, however, have argued that the Great Moderation was simply the result of sheer good luck (Stock and Watson, 2003; Justiniano and Primiceri, 2008): a general decrease in both the frequency and the intensity of exogenous (supply-side) shocks. In this sense, good policies had nothing to do with it.

But what are these good policies and good theories? According to this Great Consensus, good policies rest on what Rochon and Setterfield (2008) call 'monetary policy dominance', that is, the idea that only monetary policy is capable of regulating the business cycle, which in turn rests on the necessity for an independent central bank.

One of the immediate implications was to cast aside the role traditionally played by fiscal policy, associated with so-called Keynesian economics. In fact, strong anti-Keynesian policies, such as austerity policies and 'fiscal consolidation' policies (the idea that economic growth comes from cutting back public expenditures), became the norm. Traditional Keynesian policies were seen as inflationary, contributing to high interest rates, which in turn discourage private investment. It was a return to pre-Keynesian notions of virtuous markets; economists abandoned fiscal policy.

In this regard, Blinder (2004, p. 1) noted that: 'Virtually every contemporary discussion of stabilization policy by economists – whether it is abstract or concrete, theoretical or practical – is about monetary policy, not fiscal policy.' Farrell and Quiggin (2012, p. 21) argued along similar lines:

> The dominant approach to macroeconomic policy was based on the assumption that an independent central bank, adjusting short-term interest rates in line with a 'Taylor rule', could manage the economy in such a way as to achieve both stable inflation and reasonably steady economic growth. Active fiscal policy could not improve on this outcome, and would effectively be neutralised by offsetting adjustments to monetary policy. The 'Great Moderation' (a general reduction in the volatility of output, prices and employment beginning in the 1980s) was seen as the happy outcome of this policy framework.

These statements are simply a restatement of Lucas's (1980, p. 19) famous observation that:

> One cannot find good, under-forty economists who identify themselves or their work as 'Keynesian'. Indeed, people even take offense if referred to as 'Keynesians'. At research seminars, people don't take Keynesian theorizing seriously anymore;

> the audience starts to whisper and giggle to one another. Leading journals aren't getting Keynesian papers submitted any more.

Strangely enough, this view was well established even during the 2007–08 crisis itself. Almost two years after the beginning of the crisis, *The Economist* magazine (2009), in its rejection of Keynesian-inspired fiscal policies, went so far as to suggest that '[r]eal scientists, after all, do not leaf through Newton's 'Principia Mathematica' to solve contemporary problems in physics'. The implicit reference to Keynes's *General Theory*, written in 1936 (Keynes, 1936/1964), was obvious, and the implication was even more so: economists should not bother finding inspiration in a book that was written some eight decades earlier. And for the economics profession at large, this was essentially correct. As Kelton and Wray (2006, p. 101) point out: 'This instinctive turn towards monetary policy for stabilization represents the culmination of a long-term trend away from "Keynesian" reliance on fiscal policy.'

With respect to good theory, it was argued that macroeconomics had to be based on microeconomics; or microfoundations. This then led to the emergence of a whole new generation of highly technical models such as the DSGE model, which were said to be more sophisticated. The claim was that such models now elevated economics to the level of a true science (like chemistry and physics), and as such, like physics, contained fundamental laws of behaviour that must be 'true'. These so-called laws of behaviour rested on a representative agent's rational behaviour and their intertemporal choices (Box 2.6).

In the ironic description by Solow (2008, p. 243), this amounts to:

> macroeconomics that is deduced from a model in which a single immortal consumer–worker–owner maximizes a perfectly conventional time-additive utility

BOX 2.6

INTERTEMPORAL CHOICES

In mainstream economics, the representative agent behaves rationally, meaning that they seek to maximize their utility or satisfaction. This implies that they must also make decisions regarding the present versus the future. In other words, should an agent spend all their income today or save some in order to raise their income in the future? When taking this decision, the agent will compare the rate of interest (on savings) with their preferences for consuming today relative to the future.

> function over an infinite horizon, under perfect foresight or rational expectations, and in an institutional and technological environment that favors universal price-taking behavior.

Microfoundations and the use of DSGE models transformed economics profoundly. Up until then, Keynesians had argued that markets were imperfect and could lead to temporary breakdowns, resulting in unemployment or low economic growth that required the intervention of the government through the active use of fiscal policy. Now, it was argued, markets were rational, and as such could not be systematically wrong. The Great Moderation validated this interpretation of events, and convinced many in the economics profession that the 'improved micro-founded macroeconomics' had greatly reduced economic instability. So much so that when the first signs of the crisis appeared, as a result of the bursting of a real-estate bubble, they were dismissed as 'noise'. After all, how can there be a crisis when the model says otherwise?

For instance, in 2005, Bernanke (2005) stated that the sharp real-estate price increases observed at the time (which worried many economists) 'largely reflect strong economic fundamentals'. Eugene Fama also echoed these sentiments: 'Housing markets are less liquid, but people are very careful when they buy houses. It's typically the biggest investment they're going to make, so they look around very carefully and they compare prices. The bidding process is very detailed' (Clement, 2007). In other words, so convinced were they of the wisdom of their models, and the apparent rationality of economic agents, that they ignored the early rumblings of the financial crisis. Whatever was going on must be sound, since their model told them so.

That same year, in his Congressional testimony in March, Bernanke (2007) declared: 'At this juncture, the impact on the broader economy and financial markets of the problems in the subprime market seems likely to be contained.' On 10 January 2008, he even stated that '[t]he Federal Reserve is not currently forecasting a recession' (Bernanke, 2008). Famous last words.

Yet, there were some who were ringing the alarm bells. For instance, at the 2005 Jackson Hole conference in honour of Alan Greenspan, Raghuram Rajan (Chief Economist at the IMF from 2003 to 2007; currently Vice-Chairman at the Bank for International Settlements) presented a paper warning of unsustainable practices in financial markets (Box 2.7).

But once again, because the models could not predict a crisis, and because it was assumed that markets and agents were all rational, Rajan's warnings were

BOX 2.7

JACKSON HOLE CONFERENCES

The Jackson Hole Economic Symposium is an important annual gathering, held in late August, of central bankers, finance ministers and academics, and has been sponsored by the Federal Reserve Bank of Kansas City since 1978. Since 1981, it has been held in the town of Jackson Hole, Wyoming, USA. Each year, the conference addresses an important topic, and the discussions tend to carry weight with policy-makers. Topics that were covered in the recent past include 'Changing market structures and implications for monetary policy' (in 2018), and 'Challenges for monetary policy' (in 2019).

dismissed, and Summers called his research 'misguided' and referred to him as a 'luddite' (*The Economist*, 2010).

With the failure of Lehman Brothers in September 2008, however, it was more difficult to continue denying a financial crisis was actually happening (with early signs in the subprime market in 2006). It is at this point that the consensus unravelled, and discord set in. The consequences were enormous, among which was a return to policy discussions that included fiscal policy and Keynesian economics, or what Robert Skidelsky (2009) called, and titled his book, *The Return of the Master*; namely, the economics of John Maynard Keynes.

Rethinking DSGE models?

The 2007–08 financial crisis marked the beginning of the Great Civil War in macroeconomics, and the old divisions, which for a while were glossed over in light of the Great Moderation, quickly re-emerged. Fingers were pointed especially at the miserable failure of microfounded models. After all, despite the sophistry of the DSGE model, the simple reality is that it had failed to predict and foresee the coming of the crisis, prompting even the Queen of the United Kingdom to now famously wade publicly into the debate, and contribute to the admonishment of the economics profession.

From the crisis onward, economists were busy accusing each other of grave mistakes, and worse. For his part, Krugman (2013) pulled no punches, and attacked the quest for microfoundations: 'So the truth was that microfoundations in macroeconomics had its moment, but failed utterly at the one thing

it was sold, above all, as being able to do – namely, give a better explanation of why nominal shocks have real effects. Time, you might think, to reconsider the project.' A few years earlier, Krugman (2009a) had claimed that: 'The economics profession mistook beauty, clad in impressive-looking mathematics, for truth.'

There were others as well. Hope and Soskice (2016, p. 218) more recently described these models as 'aesthetically beautiful but mad . . . The result has increasingly been macro models of great complexity that bear little relation to reality.' This is precisely what King (2012) called the 'microfoundations delusion' (see the concluding chapter of this volume).

Bradford DeLong stated that Lucas and his followers were 'making ancient and basic analytical errors all over the place' (*The Economist*, 2009). Paul Krugman (2009b) wrote that Robert Barro (professor at Harvard University, a leader in the use of these models) is 'making truly boneheaded arguments'. Along similar lines, there is an often-referenced opinion by Solow (2003) given on 25 October 2003 on the occasion of Joseph Stiglitz's sixtieth birthday. Entitled 'Dumb and dumber macroeconomics', Solow says: 'The original impulse to look for better or more explicit micro foundations was probably reasonable . . . What emerged was not a good idea.'

Krugman even went further. In his Lionel Robbins lectures at the London School of Economics on 10 June 2009, he claimed that, in the last three decades or so, most macroeconomics was 'spectacularly useless at best, and positively harmful at worst' (quoted by *The Economist*, 2009) (Box 2.8).

Krugman was not alone in sharing this sentiment. According to Solow (2010), in a 2010 address to the United States Congress:

BOX 2.8

LIONEL ROBBINS LECTURES

Born in 1898, Lionel Robbins was a Professor of Economics at the London School of Economics for over 30 years, and Chairman of the Court of Governors. He worked closely on the creation of the Bretton Woods agreement, which in turn established the International Monetary Fund and the World Bank. He initially opposed the work of John Maynard Keynes, and in 1934 he wrote a book on the Great Depression that was largely anti-Keynesian. He came to renounce the contents of that book, in favour of the policies of Keynes. The annual lectures honour his memory.

> I do not think that the currently popular DSGE models pass the smell test. They take it for granted that the whole economy can be thought about as if it were a single, consistent person or dynasty carrying out a rationally designed, long-term plan, occasionally disturbed by unexpected shocks, but adapting to them in a rational, consistent way . . . The protagonists of this idea make a claim to respectability by asserting that it is founded on what we know about microeconomic behavior, but I think that this claim is generally phony. The advocates no doubt believe what they say, but they seem to have stopped sniffing or to have lost their sense of smell altogether.

All this adds to the often-quoted passage by Buiter (2009):

> The typical graduate macroeconomics and monetary economics training received at Anglo-American universities during the past 30 years or so, may have set back by decades serious investigations of aggregate economic behavior and economic policy-relevant understanding. It was a privately and socially costly waste of time and resources.

Nobel laureate Joseph Stiglitz (2002) once argued that economics as taught 'in America's graduate schools . . . bears testimony to a triumph of ideology over science'. But even back as early as 1991, Stiglitz along with Blanchard, Summers, Arrow, Blinder, and other prominent economists, in a joint statement of the Report of the Commission on Graduate Education in Economics, were already ringing the bell: 'Graduate programs [in economics] may be turning out a generation with too many *idiots savants*, skilled in technique but innocent of real economic issues' (Krueger et al., 1991, pp. 1044–5).

More recently, 2018 Nobel laureate Paul Romer (2016, p. 1) weighed in decisively with a familiar theme: 'For more than three decades, macroeconomics has gone backwards . . . It is sad to recognize that economists who made such important scientific contributions in the early stages of their careers followed a trajectory that took them away from science.' He then compared those who defend these unrealistic models as anti-vaxers and homeopaths, and likened the models themselves to 'Phlogiston'. He went so far as to say that it might 'masquerade as science . . . [and] might survive as entertainment'.

In the face of these accusations, one would expect resistance from those whose views were being attacked. The strategy from the defenders of these models was to protect the integrity of the DSGE model, and as such to take blame away from the DSGE model itself: the models were sound, they argued. In other words, the crisis had nothing to do with either policy or theory emanating from these microfounded models: the models reflected

sound economics, and the inability of the models to predict the crisis cannot be seen as a flaw of them. In a show of twisted logic, since the models were not built to predict crises, the fact they did not predict the crisis is proof that the models are sound.

For instance, Blanchard et al. (2010, p. 16) argued that: 'The crisis was not triggered primarily by macroeconomic policy . . . In many ways, the general policy framework should remain the same.' Along similar lines, Bernanke (2010, pp. 2–4) argued in this regard that:

> although economists have much to learn from this crisis . . . I think that calls for a radical reworking of the field go too far . . . I would argue that the recent financial crisis was more a failure of economic engineering and economic management than of what I have called economic science. The economic engineering problems were reflected in a number of structural weaknesses in our financial system . . . Shortcomings of what I have called economic science, in contrast, were for the most part less central to the crisis; indeed, although the great majority of economists did not foresee the near-collapse of the financial system, economic analysis has proven and will continue to prove critical in understanding the crisis, in developing policies to contain it, and in designing longer-term solutions to prevent its recurrence . . . In short, the financial crisis did not discredit the usefulness of economic research and analysis by any means.

Hence, according to mainstream economists, the models could do no wrong. As stated above, because their models assumed perfectly rational economic agents, the possibility that agents could be wrong was rejected. According to their models, after all, a crisis could not occur. As Krugman (2009a) wrote: 'More important was the profession's blindness to the very possibility of catastrophic failures in a market economy.'

This was recognized even by those who defended the models, albeit after the 2007–08 financial crisis. For instance, Bernanke (2010, p. 17) stated quite clearly that 'the standard models were designed for these non-crisis periods'. Others have argued along similar lines. For instance, Charles Goodhart (2009, p. 352) writes that these models 'were, by construction, fair weather models only', meaning that they could only be applied in good times. Blanchard (2014, p. 28) was perhaps the clearest: these models 'were best suited to a worldwide view in which economic fluctuations occurred but were regular, and essentially self-correcting'.

In an interview with the Federal Reserve Bank of Minneapolis, Nobel laureate Thomas Sargent (2010) – like Barro, a defender of these models – said

the same thing, albeit rather defensively: 'These models were designed to describe aggregate economic fluctuations during normal times when markets can bring borrowers and lenders together in orderly ways, not during financial crises and market breakdowns.'

In many ways, the financial crisis has put the mainstream modellers on the defensive. Yet, none of them are prepared to lay blame on the models. This reminds us of what Keynes (1936, p. 16) stated about parallel lines:

> The [neo]classical theorists resemble Euclidean geometers in a non-Euclidean world who, discovering that in experience straight lines apparently parallel often meet, rebuke the lines for not keeping straight – as the only remedy for the unfortunate collisions which are occurring. Yet, in truth, there is no remedy except to throw over the axiom of parallels and to work out a non-Euclidean geometry. Something similar is required to-day in economics.

Mainstream modellers refuse to 'rebuke' their models.

One obvious question then is: why even use these models if they have all these flaws? The answer is that the 2007–08 financial crisis has opened up legitimate questioning of these models and their performance, and some economists, both within the mainstream and outside of it, are now openly questioning these flawed models.

An alternative perspective: rejecting the models?

The increased criticism and considerable dissatisfaction aimed at the use of DSGE models led David Vines and Samuel Wills of the University of Oxford to create the Rebuilding Macroeconomic Theory project, and to publish a double issue of the *Oxford Review of Economic Policy* in Spring–Summer 2018 aimed at exploring the weaknesses of these models. Comprising 13 papers from a number of well-established economists (see Blanchard, 2018; Carlin and Soskice, 2018; Ghironi, 2018; Haldane and Turrell, 2018; Hendry and Muelbauer, 2018; Krugman, 2018; Lindé, 2018; McKibbin and Stoeckel, 2018; Reis, 2018; Stiglitz, 2018; Vines and Willis, 2018; Wren-Lewis, 2018; Wright, 2018), this double issue provides a number of criticisms that are worth considering. In a way, this project aims at reconciling the economics profession by focusing economists' investigation away from the infighting and toward the need to find common ground around the methodological issues. Yet, the title of the project itself, Rebuilding Macroeconomic Theory, is telling: it is an admission of how broken macroeconomics has become. But does rebuilding imply new views and new approaches, or rebuilding the old ones?

As stated at the beginning of this chapter, Vines and Willis (2018, p. 2) admit that it is 'no longer clear what macroeconomic theory should look like, or what to teach the next generation of students. We are still looking for the kind of constructive response to this crisis that Keynes produced in the 1930s.' Indeed, Keynes's (1936/1964) *General Theory* gave policy-makers a blueprint to recovery.

It is impossible here to review all the papers in the double issue of the *Oxford Review of Economic Policy* in Spring–Summer 2018, although that remains a valuable exercise. We will nevertheless review a few of these papers here, in order to give a general sense of the conclusions of this ambitious project. And the general conclusion, at first glance, appears rather positive; although, as we shall conclude, appearances can be deceiving. But at least, it begins on a good footing, with Vines and Willis (2018, p. 4) asking: 'is the new-Keynesian DSGE model fit for purpose? Most authors agree that the answer is no.' So far so good. Yet, this answer does not mean that the economics profession is ready to forget the infamous model, and the authors make that very clear. Indeed, Vines and Willis (2018, p. 4) are clear on what needs to be done, and this simply amounts to 'tweaking the model', or what King (2012, p. 150) claims amounts to a tug-of-war about 'my microfoundations are better than yours'.

According to Vines and Wills (2018), any attempt to 'reconstruct' macroeconomics needs to respect the following conditions:

1. incorporate financial frictions rather than assume financial intermediation is costless;
2. relax the requirement of rational expectations;
3. introduce heterogeneous agents; and
4. underpin the model – and each of these three new additions – with more appropriate microfoundations.

In other words, as the French say, '*plus ça change, plus les choses restent pareilles*': the more things change, the more they remain the same. In Vines and Wills's (2018) prescription, two arguments stand out: (1) the need to 'relax' rational expectations – and not to abandon them; and (2) the need for 'more appropriate' microfoundations – and not to abandon them. Combined, we see right away the limits of the proposed reforms. This precaution is carried through most of the contributions to the journal's symposia.

Among the most critical papers, British economist Wren-Lewis (2018, p. 55) does not believe in abandoning the conventional DSGE model, but rather

wishes to see parallel models exist alongside it: 'Just as microfoundations hegemony held back macroeconomics on that occasion, it is likely to do so again. Macroeconomics will develop more rapidly in useful directions if microfounded models and more traditional SEMs [structural equation modelling – models with many equations] work alongside each other, and are accorded equal academic respect.' Wren-Lewis (2016, p. 27) had made a similar statement two years before:

> The problem as I see it is more that in becoming the only accepted way of doing serious macroeconomic research and policy analysis, it [DSGE] (quite deliberately) crowded out other more traditional approaches that – had they persisted – might have left the discipline in a better position to understand the impact of the financial crisis. Furthermore, these alternative approaches might have encouraged the development of DSGE models in what we now know to be more fruitful directions.

A somewhat reformed Blanchard (2018, p. 51) takes a similar view. His solution is to have 'different types of macro models' co-existing at the same time (p. 44): a theory model (DSGE model) and a number of policy models (data-driven models), each with their respected strengths and weaknesses, but in the end they are equal or at least 'all needed':

> Not all models have to be explicitly microfounded. While this will sound like a plaidoyer *pro domo*, I strongly believe that *ad hoc* macro models, from various versions of the IS–LM to the Mundell–Fleming model, have an important role to play in relation to DSGE models. They can be useful upstream, before DSGE modelling, as a first cut to think about the effects of a particular distortion or a particular policy. (Blanchard, 2018, p. 48)

To be sure, Blanchard (2018) is not questioning the need to reject DSGE models: 'current DSGE models are flawed. But DSGE models can fulfil an important need in macroeconomics; that of offering a core structure around which to build and organize discussions' (p. 48), and they must be 'built on explicit micro foundations' (p. 49).

By far the most critical contributions are Krugman (2018) and Haldane and Turrell (2018). While Krugman (2018) centres his criticism around the traditional macro Keynesian model and how well it has performed over the years, Haldane and Turrell (2018) call for an interdisciplinary approach to macroeconomics, and even cite well-known post-Keynesian authors. In the end, Haldane and Turrell (2018, p. 243) still argue in favour of a parallel model: 'microfounded models are not the only kind of models that are

useful for making sense of aggregate economic fluctuations. A more diverse approach to macroeconomic modelling may be beneficial when making sense of the economy and when setting policy to shape the economy.'

The post-Keynesian reply

From the post-Keynesian perspective, the crisis in macroeconomics reflects its deep misunderstanding of how economies work. Models are as relevant as their assumptions. In this sense, rational agents, absence of true uncertainty, microfoundations of any kind, and equilibrium analysis are simply not realistic assumptions of the world in which we live. It therefore comes as no surprise that post-Keynesians and heterodox authors reject DSGE models: they simply do not do a good job at modelling the real world and therefore offer very little in terms of appropriate policies. The fact that many in the profession still insist on keeping the DSGE approach around is in itself a problem.

For post-Keynesian economists, as the chapters in this book will show, the real world is characterized by uncertainty about the future, and it is misplaced to use models that assume rational behaviour, or knowledge about the future. Post-Keynesian economists believe that macroeconomics provides the proper foundations for microeconomics, because our society is rooted in institutions and social classes, and that such a view of the world cannot be reduced to a single 'representative' agent. The concluding chapter of this book, by John King, explores this theme in great detail.

As for the civil war in economics, there is considerable resistance to change. The profession stubbornly resists abandoning its flawed models because of the belief in TINA. The chapters in this book will show that a coherent alternative does exist.

Conclusion

In this chapter, we have argued that while the 2007–08 financial crisis has led many within the mainstream to rethink some of their methodology, models and theories, the reflection has obviously not gone far enough. So far, the exercise seems to be limited to a lukewarm criticism, a certain dose of *mea culpa*, and to quickly move on to the same things. In this sense, the mainstream dissent is not really dissent, or is 'dissent-light': let us keep our model but let us find 'better' microfoundations. This is precisely Stiglitz's (2018, p. 70) assessment as well: 'at the heart of the failure were the wrong microfoundations, which failed to incorporate key aspects of economic behaviour,

e.g. incorporating insights from information economics and behavioural economics.'

The conclusion, as Lavoie (2018, p. 3) argues, is that 'several mainstream economists . . . believe that mainstream macroeconomics can be easily repaired from within, or that the required modifications already existed before the crisis but were ignored'.

The economics profession is not only unwilling but also unable to go any further, largely because it does not see any alternative: 'what we have may not be perfect, but it's all we've got'. For instance, Lindé (2018, p. 269) argues that 'DSGE models currently have few contenders to replace them as core models in the policy process'. According to Krugman (2018, p. 157), 'there hasn't been a big new idea, let alone one that has taken the profession by storm'. This much has been confirmed, in an interview on Bloomberg (18 November 2016), by Chari from the University of Minnesota: '[b]urning down the edifice, and saying we'll figure out what we'll build on its foundations later, just does not seem like a constructive way to proceed' (Bloomberg, 2016).

In the end, one can imagine that in the absence of an alternative or the next big idea, the economics profession will attempt to repeat history, by adding bits and pieces to the existing model until it better reflects their perception of a microfounded world, that is, a new new-neoclassical synthesis. For instance, Akerlof and Shiller (2009, p. 268) argue that:

> It is necessary to incorporate animal spirits into macroeconomic theory in order to know how the economy really works. In this respect the macroeconomics of the past thirty years has gone in the wrong direction. In their attempt to clean up macroeconomics and make it more scientific, the standard macroeconomists have imposed a research structure and discipline by focusing on how the economy would behave if people had only economic motives and they were also fully rational.

Blanchard (2009, p. 214) observed in this respect that 'the new tools developed by the new-classicals came to dominate. The facts emphasized by the new-Keynesians forced imperfections back in the benchmark model. A largely common vision has emerged.' He later on returned to this idea stating that:

> As a result of the crisis, a hundred intellectual flowers are blooming. Some are very old flowers: Hyman Minsky's financial instability hypothesis, Kaldorian models of growth and inequality. Some propositions that would have been considered

BOX 2.9

HYSTERESIS

Hysteresis is an old concept which suggests that the current state of a system depends on its history. In economics, this would suggest that the concept of equilibrium cannot exist independently of the state of the economy. In mainstream economics, for instance, an equilibrium exists independently of the forces of supply and demand; a position towards which the economy gravitates. But post-Keynesians argue that this 'equilibrium' depends on the state of the economy; it is not independent of it. For instance, if the economy is depressed for some time, this in turn will lower the long-run rate of economic growth. In other words, the long-run rate of growth depends on the rate of economic growth in the short run.

> anathema in the past are being proposed by 'serious' economists. For example, monetary financing of the fiscal deficit. Some fundamental assumptions are being challenged, for example the clean separation between cycles and trends. Hysteresis is making a comeback. This is all for the best. (International Monetary Fund, 2015)

For the record, Minsky and Kaldor are well-known post-Keynesian economists, and the concept of hysteresis is a core post-Keynesian principle (Box 2.9).

So Blanchard is suggesting adding to the mainstream model some post-Keynesian concepts. This certainly will not be enough to provide sound macroeconomic analyses and policy decisions.

The only glimmer of hope seems to be coming from outside of academia. For instance, in 2011 the Independent Evaluation Office of the IMF acknowledged that its research had failed (or rather, 'fell short') to deliver on its core mission of alerting countries to their economic vulnerabilities. In its own analysis, the Independent Evaluation Office attributed this IMF failure to a 'high degree of groupthink; intellectual capture; and a general mindset that a major financial crisis in large advanced economies was unlikely' (International Monetary Fund, 2011, p. 17), because the IMF economists had blindly adopted a theory that could not even consider the possibility of such a crisis. To remedy this failure, the lack of dissenting views within the IMF is no longer acceptable, and the IMF should 'actively seek alternative or dissenting views' (International Monetary Fund, 2011, p. 21). A few years later, in a now-famous document, the IMF openly asked whether mainstream economics or neoliberalism had been 'oversold' (Ostry et al., 2016).

This revolution is spreading. Now the European Central Bank (ECB) is casting serious doubt on the 'unrealistic' DSGE model: at a speech in Frankfurt in September 2017, Vítor Constâncio, at the time Vice-President of the ECB, listed a number of reasons why the standard model is no longer relevant for policy research (Constâncio, 2017). More recently, Andy Haldane, Chief Economist at the Bank of England, explained that economists often lack creativity, and tried to inspire them by inviting guest speakers such as poets, Olympians, ballerinas and percussionists, so that economists can fully experience the 'wellspring of creativity' (Haldane, 2018).

Let us hope that the economics profession will listen to these important calls for urgent reforms before the next financial crisis erupts. The world economy and its population are running out of both time and hope.

NOTE

1 To be fair, some versions of the DSGE model did include financial frictions, such as in Bernanke et al. (1999), albeit these frictions are not a dominant feature of this model.

REFERENCES

Akerlof, G.A. and R.J. Shiller (2009), *Animal Spirits*, Princeton, NJ: Princeton University Press.

Bernanke, B. (2005), 'The economic outlook', Statement to the Council of Economic Advisers, 20 October, https://georgewbush-whitehouse.archives.gov/cea/econ-outlook20051020.html.

Bernanke, B.S. (2007), 'The economic outlook', Testimony before the Joint Economic Committee of the US Congress, Board of Governors of the Federal Reserve System, 28 March, accessed 27 August 2018 at www.federalreserve.gov/newsevents/testimony/bernanke20070328a.htm.

Bernanke, B. (2008), 'Fed ready to cut interest rates again', public statement, accessed 24 August 2018 at www.nbcnews.com/id/22592939/ns/business-stocks_and_economy/t/bernanke-fed-ready-cut-interest-rates-again.

Bernanke, B.S. (2010), 'On the implications of the financial crisis for economics', remarks at the conference co-sponsored by the Bendheim Center for Finance and the Center for Economic Policy Studies, Princeton University, 24 September, accessed 24 August 2018 at www.federalreserve.gov/newsevents/speech/bernanke20100924a.htm.

Blanchard, O.J. (2009), 'The state of macro', *Annual Review of Economics*, **1**, 209–28.

Blanchard, O.J. (2014), 'Where danger lurks', *Finance and Development*, **51** (3), 28–31.

Blanchard, O.J. (2018), 'On the future of macroeconomic models', *Oxford Review of Economic Policy*, **34** (1–2), 43–54.

Blanchard, O.J., G. Dell'Ariccia and P. Mauro (2010), 'Rethinking macroeconomic policy', International Monetary Fund Staff Position Note, No. 10/03.

Blinder, A.S. (2004), 'The case against the case against discretionary fiscal policy', Center for Economic Policy Studies Working Paper, No. 100.

Bloomberg (2016), 'The rebel economist who blew up macroeconomics', accessed 16 November 2017 at www.bloomberg.com/news/articles/2016-11-18/blah-blah-blah-a-renowned-economist-sums-up-the-state-of-macro.

Bloomberg (2020), 'Paul Krugman is pretty upbeat about the economy right now', 27 May, accessed 4 July 2020 at www.bloomberg.com/opinion/articles/2020-05-27/paul-krugman-is-pretty-upbeat-about-coronavirus-economic-recovery.

Buiter, W. (2009), 'The unfortunate uselessness of most "state of the art" academic monetary economics', Munich Personal RePEc Archive Paper, No. 58407, accessed 24 August 2018 at https://mpra.ub.uni-muenchen.de/58407/1/MPRA_paper_58407.pdf.

Carlin, W. and D. Soskice (2018), 'Stagnant productivity and low unemployment: stuck in a Keynesian equilibrium', *Oxford Review of Economic Policy*, **34** (1–2), 169–94.

Chari, V.V. and P.J. Kehoe (2006), 'Modern macroeconomics in practice: how theory is shaping policy', *Journal of Economic Perspectives*, **20** (4), 3–28.

Clement, D. (2007), 'Interview with Eugene Fama', *Region*, Federal Reserve Bank of Minneapolis, December, accessed 27 August 2018 at www.minneapolisfed.org/publications/the-region/interview-with-eugene-fama.

Constâncio, V. (2017), 'Developing models for policy analysis in central banks', opening speech at the Annual Research Conference, Frankfurt am Main, 25 September, accessed 27 August 2018 at www.ecb.europa.eu/press/key/date/2017/html/ecb.sp170925.en.html.

Coric, B. (2011), 'The sources of the Great Moderation: a survey', mimeo.

Davis, S. and J. Kahn (2008), 'Interpreting the Great Moderation: changes in the volatility of economic activity at the macro and micro levels', *Journal of Economic Perspectives*, **22** (4), 155–80.

The Economist (2009), 'The other-worldly philosophers', 16 July, accessed 12 August 2018 at www.economist.com/briefing/2009/07/16/the-other-worldly-philosophers.

The Economist (2010), 'Larry Summers: neo-Keynesian aristocrat', 8 October, accessed 24 August 2018 at www.economist.com/democracy-in-america/2010/10/08/larry-summers-neo-keynesian-aristocrat.

Farrell, H. and J. Quiggin (2012), 'Consensus, dissensus and economic ideas: the rise and fall of Keynesianism during the economic crisis', mimeo, accessed 24 August 2018 at www.henryfarrell.net/Keynes.pdf.

Ghironi, F. (2018), 'Macro needs micro', *Oxford Review of Economic Policy*, **34** (1–2), 195–218.

Goodhart, C. (2009), 'The continuing muddles of monetary theory: a steadfast refusal to face the facts', in E. Hein, T. Niechoj and E. Stockhammer (eds), *Macroeconomic Policies on Shaky Foundations: Whither Mainstream Economics?*, Marburg: Metropolis, pp. 351–69.

Haldane, A.G. (2018), 'What Grayson Perry can teach the Bank of England', *Prospect Magazine*, 18 April, accessed 24 August 2018 at www.prospectmagazine.co.uk/magazine/new-thinking-at-the-old-lady-the-bank-of-englands-unlikely-new-advisers.

Haldane, A.G. and A.E. Turrell (2018), 'An interdisciplinary model for macroeconomics', *Oxford Review of Economic Policy*, **34** (1–2), 219–51.

Hall, R. (1976), 'Notes on the current state of empirical macroeconomics', accessed 4 July 2020 at https://web.stanford.edu/~rehall/Notes%20Current%20State%20Empirical%201976.pdf.

Hendry, H. and J.N.J. Muellbauer (2018), 'The future of macroeconomics: macro theory and models at the Bank of England', *Oxford Review of Economic Policy*, **34** (1–2), 287–328.

Hope, D. and D. Soskice (2016), 'Growth models, varieties of capitalism and macroeconomics', *Politics and Society*, **44** (2), 209-26.

International Monetary Fund (2011), 'IMF performance in the run-up to the financial and economic crisis', Independent Evaluation Office, accessed 28 August 2018 at www.ieo-imf.org/ieo/files/completedevaluations/Crisis-%20Main%20Report%20(without%20Moises%20Signature).pdf.

International Monetary Fund (2015), 'Blanchard: looking forward, looking back', interview with Olivier Blanchard, 31 August, accessed 27 August 2018 at www.imf.org/en/News/Articles/2015/09/28/04/53/sores083115a.

Justiniano, A. and G.E. Primiceri (2008), 'The time-varying volatility of macroeconomic fluctuations', *American Economic Review*, **98** (3), 604–41.

Kelton, S. and L.R. Wray (2006), 'What a long, strange trip it's been: can we muddle through without fiscal policy?', in C. Gnos and L.-P. Rochon (eds), *Post-Keynesian Principles of Economic Policy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 101–19.

Keynes, J.K. (1936/1964), *The General Theory of Employment, Interest and Money*, London: Harcourt Brace.

King, J.E. (2012), *The Microfoundations Delusion: Metaphor and Dogma in the History of Macroeconomics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Krueger, A.O., K.J. Arrow, O.J. Blanchard, C. Goldin, E.E. Leamer, et al. (1991), 'Report of the Commission on graduate education in economics', *Journal of Economic Literature*, **29** (3), 1035–53.

Krugman, P. (2009a), 'How did economists get it so wrong?', *New York Times Magazine*, 2 September, accessed 27 August 2018 at www.nytimes.com/2009/09/06/magazine/06Economic-t.html.

Krugman, P. (2009b), 'War and non-remembrance', *New York Times Blogs*, 22 January, accessed 28 August 2018 at https://krugman.blogs.nytimes.com/2009/01/22/war-and-non-remembrance/.

Krugman, P. (2013), 'Microfoundations and the parting of the waters', *New York Times Blogs*, 20 December, accessed 24 August 2018 at https://krugman.blogs.nytimes.com/2013/12/20/microfoundations-and-the-parting-of-the-waters/.

Krugman, P. (2018), 'Good enough for government work? Macroeconomics since the crisis', *Oxford Review of Economic Policy*, **34** (1–2), 156–68.

Lavoie, M. (2018), 'Rethinking macroeconomic theory before the next crisis', *Review of Keynesian Economics*, **6** (1), 1–21.

Lindé, J. (2018), 'DSGE models: still useful in policy analysis?', *Oxford Review of Economic Policy*, **34** (1–2), 269–86.

Lucas, R. (1980), 'The death of Keynesian economics', *Issues and Ideas*, Winter, 18–19.

Lucas, R. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics*, **22** (1), 3–42.

Lucas, R. (2003), 'Macroeconomic priorities', *American Economic Review*, **93** (1), 1–14.

Mankiw, G. (2006), 'The macroeconomist as scientist and engineer', *Journal of Economic Perspectives*, **20** (4), 29–46.

McKibbin, W. and A. Stoeckel (2018), 'Modelling a complex world: improving macro-models', *Oxford Review of Economic Policy*, **34** (1–2), 329–47.

Ostry, J.D., P. Loungani and D. Furceri (2016), 'Neoliberalism: oversold?', *Finance and Development*, **53** (2), 38–41.

Reis, R. (2018), 'Is something really wrong with macroeconomics?', *Oxford Review of Economic Policy*, **34** (1–2), 132–55.

Rochon, L.-P. and M. Setterfield (2008), 'The political economy of interest rate setting, inflation, and income distribution', *International Journal of Political Economy*, **37** (2), 2–25.

Romer, P. (2016), 'The trouble with macroeconomics', mimeo, accessed 24 August 2018 at https://paulromer.net/wp-content/uploads/2016/09/WP-Trouble.pdf.

Sargent, T. (2010), 'Modern macroeconomics under attack', *Region*, **24** (3), 28–39, accessed 27 August 2018 at https://www.minneapolisfed.org/~/media/files/pubs/region/10-09/sargent.pdf.

Skidelsky, R. (2009), *Keynes: The Return of the Master*, New York: Public Affairs.

Solow, R. (2003), 'Dumb and dumber in macroeconomics', address at Joseph Stiglitz's 60th birthday conference, accessed 27 August 2018 at http://economistsview.typepad.com/economistsview/2009/08/solow-dumb-and-dumber-in-macroeconomics.html.

Solow, R. (2008), 'The state of macroeconomics', *Journal of Economic Perspectives*, **22** (1), 243–9.
Solow, R. (2010), 'Building a science of economics for the real world', House Committee on Science and Technology, Subcommittee on Investigations and Oversight, 20 July.
Sraffa, P. (1926), 'The laws of returns under competitive conditions', *Economic Journal*, **36** (144), 535–50.
Sraffa, P. (1930), 'Increasing returns and the representative firm: a criticism', *Economic Journal*, **40** (157), 89–92.
Sraffa, P. (1951), 'Introduction', in D. Ricardo, *The Works and Correspondence of David Ricardo, Vol. I*. edited by Piero Sraffa and Maurice Dobb, Cambridge: Cambridge University Press, pp. xiii–lxii.
Sraffa, P. (1960), *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory*, Cambridge: Cambridge University Press.
Stiglitz, J. (2002), 'There is no invisible hand', *Guardian*, 20 December, accessed 27 August 2018 at www.theguardian.com/education/2002/dec/20/highereducation.uk1.
Stiglitz, J. (2018), 'Where modern macroeconomics went wrong', *Oxford Review of Economic Policy*, **34** (1–2), 70–106.
Stock, J.H. and M.W. Watson (2003), 'Has the business cycle changed and why? Evidence and explanations', in M. Gertler and K. Rogoff (eds), *NBER Macroeconomics Annual 2002*, Boston, MA: MIT Press, pp. 159–218.
Vines, D. and S. Wills (2018), 'The Rebuilding Macroeconomic Theory project: an analytical assessment', *Oxford Review of Economic Policy*, **34** (1–2), 1–42.
Weber, S. (1997), 'The end of the business cycle?', *Foreign Affairs*, **76** (4), 65–82.
Wren-Lewis, S. (2016), 'Unravelling the new classical counter revolution', *Review of Keynesian Economics*, **4** (1), 20–35.
Wren-Lewis, S. (2018), 'Ending the microfoundations hegemony', *Oxford Review of Economic Policy*, **34** (1–2), 55–69.
Wright, R. (2018), 'On the future of macroeconomics: a New Monetarist perspective', *Oxford Review of Economic Policy*, **34** (1–2), 107–31.

A PORTRAIT OF PIERO SRAFFA (1898–1983)

by Gustavo Behring

Piero Sraffa was born in Turin, Italy on 5 August 1898 and died in 1983 in Cambridge, United Kingdom. He studied at the University of Turin from 1916 until 1920. Sraffa's undergraduate dissertation and early writings concerned monetary theory and banking, especially with regard to Italy.

His first major theoretical contribution was his 1926 paper (a shorter, revised version of the original 1925 paper in Italian), entitled 'The laws of return under competitive conditions', which criticizes Marshall's theory of costs of production. This work drew the attention of Keynes, which prompted him to invite Sraffa for a lecturing position at Cambridge, where Sraffa would meet and befriend economists such as Maurice Dobb, Joan Robinson, Nicholas Kaldor and Richard Kahn, among others, and spend the rest of his life.

In his 1926 paper, Sraffa showed that both diminishing and increasing costs are not compatible with Marshall's long-period supply curve in partial equilibrium under perfect competition. Decreasing costs due to increasing returns would cause falling marginal costs and the first firm to expand would become a monopolist, so that it was inconsistent with perfect competition. Increasing costs could be the result of decreasing returns or of an increase in the price of the factors of production used in the particular sector. But diminishing returns required a given stock of capital and were thus incompatible with long-period equilibrium conditions in which the size of capital stock of the sector has time to adjust to demand. Increasing prices of production factors would also increase to different degrees the costs and prices of other sectors and therefore were incompatible with partial equilibrium analyses of an isolated sector. The only consistent assumption was constant returns to scale.

Sraffa concluded that the only feasible way out of this inconsistency was to abandon the assumption of perfect competition. This route, however, was quickly abandoned by Sraffa, who in 1930 argued that Marshall's theory should simply be discarded, even though it greatly influenced imperfect competition theories in the 1930s, particularly by Joan Robinson and Edward Chamberlin.

Around 1927–28, Sraffa started a deeper investigation of the very concept of costs of production. He was convinced that the 'subjective' marginalist approach to value and distribution, based on presumed, but not observable, assumptions of factor substitution and consumer preferences, should give way to an 'objective' theory of value and distribution based on material costs, following the older classical approach of Petty, Smith, Ricardo (and Marx).

In 1930, Sraffa was imbued with the task of editing the Royal Economic Society's edition of *The Complete Works and Correspondence of David Ricardo* that came out in 1951. Apart from it being an outstanding scholarly endeavour for its meticulous care with the source material, the collection also presented Sraffa's new interpretation of Ricardo's theory of profits. This interpretation is explained in the introduction of the first volume, where Ricardo's theory of profits is described through what the literature nowadays calls the 'corn model', which was the basis, prior to the publication of his *Principles,*

➡

for Ricardo's *Essay on Profits*. According to Sraffa's interpretation, Ricardo maintained that the general rate of profits of the whole economy is determined in agriculture, where both the capital advanced and the product are assumed to be made up from the same commodity, namely corn. This allowed Ricardo to determine the rate of profits as a material rate only in terms of the physical cost of producing corn by means of corn. This result was subsequently generalized in Ricardo's *Principles*, to account for a heterogeneous set of commodities present in the surplus and means of production by measuring both magnitudes in terms of labour.

The parallel theoretical research programme along classical lines that had started in the late 1920s, together with the strong impact of Sraffa's new interpretation of Ricardo, culminated in the publication of Sraffa's major work, his 1960 book entitled *Production of Commodities by Means of Commodities* (*PCMC*). The aim of his book was twofold. The first was to provide a rehabilitation of the classical theory of value and distribution based on the surplus approach. *PCMC* offered a rigorous reconstruction of the general logical structure of the classical theory and provided an analytical solution, within this framework, to the determination of distribution and relative prices for given levels of output, available methods of production, and one distributive variable (either the real wage or the rate of profits). Under assumptions that were more exact and general than those of Ricardo and Marx, Sraffa demonstrates that, in spite of the rather complicated relation between changes in income distribution and relative prices, there is indeed an inverse relation between the real wage and the general rate of profits. In terms of his special construct of the standard commodity that served as an 'invariable standard of value' (pursued by Ricardo until his death), Sraffa was able to express this general negative relation in purely physical terms, as Ricardo and Marx had sought to do.

The second aim of *PCMC*, alluded to in the subtitle (*Prelude to a Critique of Economic Theory*), was to show that the complex changes in relative prices of sets of capital goods that come from the technological interdependence of the economy causes serious (and insurmountable) logical problems to the concepts of real capital endowment and the marginal product of capital (and also of the other factors of production) and therefore lays the ground for an internal critique of the marginal theory of long-period general equilibrium based on the notion of factor substitution. This critique has been further developed since then (under Sraffa's discrete supervision until the early 1970s) by many of his followers – most notably by Pierangelo Garegnani – in the so-called Cambridge capital controversies.

EXAM QUESTIONS

True or false questions

1. The COVID-19 crisis showed the degree to which governments are severally constrained in their ability to spend.
2. Relying on their models, mainstream economists are unable to propose relevant policies to get us out of recession.

3. The Great Moderation put an end to the economics of recessions.
4. The DSGE model is the only viable model for economic analysis.
5. The economics profession is ready to abandon the DSGE model.
6. Mainstream economics relies on microfoundations.
7. According to the mainstream, fiscal austerity during a recession aggravates the latter.
8. Post-Keynesian economics rejects any kinds of microfoundations.
9. The Great Moderation was the result of good economic theory and policy-making.
10. The COVID-19 crisis can induce policy-makers to adopt Keynesian policies to kick-start and support economic growth.

Multiple choice questions

1. TINA means:
   a) today is not happening;
   b) there is no alternative;
   c) toward infinite negative assumptions;
   d) none of the above.
2. 'Monetary policy dominance' implies:
   a) a mixed use of both fiscal and monetary policies;
   b) the rejection of monetary policy as a dominant policy tool;
   c) monetary policy should be the dominant policy tool;
   d) fiscal policy should be the dominant policy tool.
3. Intertemporal choices imply that households must:
   a) decide on whether they should work now versus later in life;
   b) decide whether they should save today or later;
   c) decide whether workers should receive wages today or at a later date;
   d) all of the above.
4. The Jackson Hole conference is:
   a) an annual review of papers published in leading journals;
   b) an annual address by the Governor of the Federal Reserve to the US Congress;
   c) an annual gathering of central bankers, policy-makers and economists;
   d) a student-led annual conference on ecological topics.
5. *The Return of the Master* refers to:
   a) an annual meeting of economists where fiscal policy is the main focus of discussion;
   b) an annual meeting of economists where monetary policy is the main focus of discussion;
   c) a popular documentary in which central bankers discuss the economics of Alan Greenspan;
   d) the title of a book by Robert Skidelsky.
6. Fiscal austerity is:
   a) an economic policy inspired by Keynes's writings;
   b) an economic policy inspired by the mainstream;
   c) an economic policy that has been providing good results in any case;
   d) an economic policy that has never been put into practice.
7. The Great Moderation is:
   a) a period where the rates of economic growth, inflation and unemployment were low and stable;
   b) a period where the debate among economists led to a large consensus on both theoretical and empirical grounds;

c) a period where foreign trade was lower across the global economy;
d) a period where central bankers were cooperating on a large scale.

8. Microfoundations are:
a) necessary for any kind of macroeconomic analysis;
b) useful but not necessary for any kind of macroeconomic analysis;
c) unavoidable for empirical research in macroeconomics;
d) essentially wrong for any research work in macroeconomics.

9. DSGE models are:
a) necessary for any kind of macroeconomic analysis;
b) useful but not necessary for any kind of macroeconomic analysis;
c) unavoidable for empirical research in macroeconomics;
d) essentially wrong for any research work in macroeconomics.

10. The current state of the art in macroeconomics considers DSGE models as:
a) necessary for any kind of macroeconomic analysis;
b) useful but not necessary for any kind of macroeconomic analysis;
c) unavoidable for empirical research in macroeconomics;
d) essentially wrong for any research work in macroeconomics.

3

The history of economic theories

Heinrich Bortis

OVERVIEW

- This chapter presents the historical development of two fundamentally important strands of thought in the history of economic theories, namely economics and political economy.
- It deals with the basic properties of economics: the rational behaviour of economic agents on the marketplace is supposed to be coordinated by markets such that all the resources, labour most importantly, are fully employed.
- It puts to the fore the essential features of political economy, which aims at understanding the functioning of the socioeconomic system, representing a monetary production economy.
- It emphasizes the political economy strand. François Quesnay's fundamental *tableau économique*, the theoretical founding piece of political economy, is carefully presented and its link with modern classical–Keynesian political economy established. Here the macroeconomic price equation rests on the classical labour value and surplus principles. The quantity equation (the supermultiplier) embodies Keynes's principle of effective demand. A macroeconomic equilibrium with permanent involuntary unemployment is thus possible.
- The chapter also asks which of the two approaches – economics or political economy – is more plausible. It suggests that, for both theoretical and historical reasons, modern classical–Keynesian political economy picturing a monetary production economy is definitely superior to neoclassical economics, dealing with a market economy.

KEYWORDS

- **Labour value principle:** Elaborated by David Ricardo, this states that the value of a product is, in principle, governed by the direct and indirect labour time socially necessary to produce it. Direct labour time is expended on producing the final product (say shoes), indirect labour on producing the means of production: used up fixed capital (machinery), intermediate products (leather) and basic products (electricity).
- **Marginal principle:** Of fundamental importance in economics, this concept came into being in the course of the Marginalist Revolution. The marginal principle appears in various shapes, namely marginal utility, marginal costs and marginal productivity. Marginal utility represents the additional utility occurring in response to an extra unit of some good consumed. Marginal costs are the supplementary costs arising if an additional unit of some good is produced. Marginal productivity indicates the output increase in response to employing an extra unit of some factor of production (labour, land or capital).
- **Principle of effective demand:** Established by John Maynard Keynes (1936) in his **General Theory**, this principle is based on the macroeconomic equilibrium condition 'saving = investment' and is expressed through the investment multiplier. This principle states that economic activity is, basically, not governed by supply factors (available resources) but by demand factors: public and private consumption and investment most importantly. A lack of effective demand will result in an underemployment equilibrium, implying involuntary unemployment.
- **Say's Law:** The French economist Jean Baptiste Say (1767–1832) claimed that supply creates its own demand. To produce leads to selling goods against money, which is, in turn, used to buy other goods. Hence, general overproduction and involuntary unemployment are both impossible. This supply-sided vision characterizes neoclassical mainstream economics until now.
- **Supermultiplier:** This is an elaborated form of Keynes's investment multiplier. The economy is set into motion by autonomous variables (government expenditures and exports). This brings about a cumulative process of production of both consumption and investment goods. The level of consumption is governed by the spending power of the population, depending, in turn, on the distribution of income. A more equitable distribution of income enhances the spending power of the population and, as such, is positively linked to the level of employment.
- **Surplus principle:** It was David Ricardo who neatly defined the social surplus over socially necessary (natural) wages within the framework of his corn model. Stated in modern terms, the surplus principle says that when prices are given, part of the national income accrues to the producers in the form of socially necessary wages. The remainder of the national income represents the social surplus: surplus wages, profits, land rents and labour rents due to privileges and power positions.

Why are these topics important?

Economic theories aim at the systematic explanation of economic phenomena: value and prices, income distribution, employment levels, growth and business cycles, money and finance, and international trade. These phenomena are interrelated and may be explained in very different ways, according to the theoretical approach one takes, such as classical, neoclassical, Marxian or Keynesian. For instance, for a neoclassical economist, income distribution is a market problem, whereas it is an issue of power for a Marxian political economist.

The existence of widely differing economic theories provides the basic reason why the study of the history of economic theories is required. John Maynard Keynes (1926) suggested, notably, that dealing with differing or even contradictory theories of value, distribution, employment and money leads one to independent and open-minded thinking, that is, to the 'emancipation of the mind'. This should enable the student of the history of economic theories to distil the most plausible theoretical principles, on which basis fruitful analyses and sound policy suggestions may then be founded. Hence, a socially useful teaching of economic theory must be based on the history of economic theories, as this chapter will show.

Two broad groups of economic theories

To bring fundamental issues in economic theory into the open, two broad groups of theories – namely, economics and political economy – are considered here. David Ricardo's (1821/1951) *Principles of Political Economy* and Alfred Marshall's (1890/1920) *Principles of Economics* illustrate the relevance of this distinction. Throughout this book, other chapters will refer to orthodox or mainstream views in a similar way as this chapter refers to economics, and to heterodox ideas in the way this chapter refers to political economy.

The economics approach

In economics, the great problems – value and prices, distribution and employment – are market issues, meaning that their determination is entirely from within their respective markets. The methodological approach in this regard is basically microeconomic: there are utility-maximizing households and profit-maximizing firms, whose behaviour is deemed 'rational' and is coordinated by markets in a socially meaningful way such that a social optimum prevails, namely a 'Pareto optimum', defined as a situation where no one can become better off unless someone else is made worse off. In this sense, the macroeconomic general equilibrium is derived from micro-

economic rational behaviour through the perfect functioning of markets. Only relative prices (the price of one good relative to another) are determined in neoclassical–Walrasian theory. In this theory, prices are such that the quantities demanded equal the quantities supplied on all markets, which means that an overall market equilibrium is established. Hence, the economy considered here is, essentially, an exchange economy: goods exchange for goods, and money is neutral. This implies that doubling the quantity of money doubles absolute prices (the price level of all goods), leaving relative prices and relative quantities unchanged. Schumpeter called this a 'real economy'. Moreover, money is said to be exogenous, meaning it is controlled by the central bank (see Chapter 4 for a full discussion). The distribution of income and the determination of the employment level are both market issues. Specifically, the law of supply and demand brings about full employment in competitive conditions. All those willing to work can find a job at the ruling market wage rate. The neoclassical theory gained prominence and was rendered operable by Keynes's professor, Alfred Marshall, and his supply and demand scheme, which he applied to goods and factor markets.

The political economy approach

The political economy approach stands in sharp contrast to the economics approach. In the political economy approach, the social and circular process of production, which produces a social surplus over socially necessary wages, stands at the centre of things. The fundamental prices of goods are the prices of production based on the cost of producing these goods, not market prices determined by supply and demand. Labour values form the essence of these prices of production and summarize the essential features of the social and circular process of production. As discussed in Chapter 1, social power relations positively govern the distribution of income; in a normative perspective, income distribution becomes an issue of distributive justice, constituting the heart of social ethics. Several problems arise in relation to distributive justice. First, there is the determination of wage structures within enterprises and industries based on the evaluation of workplaces. Second, trade unions have the task of broadly fixing wage differentials between industries. And, third, socially appropriate profit rates on invested capital entering price calculation have to be fixed within firms and industries. The implementation of these simple principles is, evidently, of immense complexity.

As for the level of employment, it is governed by effective demand. Unemployment is essentially involuntary: workers do not choose not to work, but do not work because no one wants to hire them. Money and finance play an essential role. Money circulating in the real sector enables the production

and circulation of goods and services, which are always exchanged against money, not against other goods. Hence, in the real sector, money always has a real-value equivalent, and money is only a representative of value. However, once money leaves the real sector to enter the financial sector, it becomes a store of value in the form of monetary wealth; as such, money has no real equivalent, but in the course of portfolio diversification looks for profitable investment in existing real and financial assets. If wealth accumulation is excessive, damaging interactions between the real sector and the financial sector may occur. Given the crucial importance of money, political economy, above all of the classical–Keynesian type, is inherently a monetary theory of production (see Chapter 4 for a full discussion).

Political economy is essentially about the socioeconomic system, made up of a material basis (that is, the economy with the social process of production at the centre), and an institutional superstructure, made up of social, legal, political and cultural institutions, which are given or evolve but slowly (Bortis, 1997, pp. 89–95). The socioeconomic system has laws of its own; most important is the principle of effective demand. Contrary to the economics perspective, there may be strong contradictions between the functioning of the system as a whole and the behaviour of economic agents. Keynes's paradox of thrift is most prominent: if all individuals save more, consumption and the volume of investment both decrease and the economy may be precipitated into a slump. This is in stark contrast to the economics approach, where more savings imply more investment and less unemployment.

In the next section, we turn to the history of economic theories. We start by sketching the historical development of economics. Subsequently, we outline the making of political economy. Finally, both strands of thought are assessed to uncover the more plausible approach fit for policy purposes.

The history of economics

Adam Smith is the founder of economics. His theory of value and distribution is dominated by the 'adding-up theorem': all prices of final goods consist of a wage, profit and rent component, meaning that the price of goods is obtained by adding up wages, profits and rents, which are the prices of the factors of production called labour, capital and land.

Adam Smith

Smith's socioeconomic system is grounded on the principle of propriety, a combination of fellow feeling and self-interest (Smith, 1759/1976a).

Contrary to modern neoclassical economics, this principle implies that ethics is on the marketplace, which means that the fundamental prices are social and, as such, fair and natural prices (Smith, 1776/1976b). For example, the money wage rate has to be such as to ensure a decent life for workers. And the price of agricultural products would have to be at a sufficiently high level to bring agricultural incomes broadly in line with incomes in industry.

However, in the short run, market prices could fluctuate around their natural level. Hence, there are two prices in Smith: a natural price, determined by the permanently acting socio-ethical principle of propriety; and a market price, which brings about deviations from the natural price, owing, for example, to extraordinary climatic conditions temporarily affecting the output of agricultural goods, to a natural calamity, or to imperfect foresight of producers who have over- or underestimated future demand. However, there is a continuous tendency for market prices to get nearer to their natural level: if the realized profit rate exceeds the natural profit rate, entrepreneurs expand productive capacities, and vice versa, thus narrowing the gap between market and natural prices. Hence, the natural price is a kind of centre of gravitation attracting market prices.

Natural prices also imply natural wages and natural profit rates, as well as natural rents on land. Hence, if all individuals acted according to the principle of propriety, inherent in human nature, a harmonious economy and society would come into being. In such an economy, full employment would naturally prevail, because in Smith's view economic activity is entirely supply-determined. Hence, the available quantities of production factors – labour, land and capital – determine the extent of economic activity, not effective demand, that is, the demand for goods and services in terms of money by consumers, producers and the state. This implies that saving is always invested. Given this, Smith explicitly speaks of the virtue of higher saving, which increases investment, hence economic growth, and as the final result enhances the wealth of a nation (Smith, 1776/1976b).

Adam Smith's theory of socio-ethical natural prices did not gain prominence. Instead, the directly visible market prices moved to the fore, giving rise to a vaguely formulated demand-and-supply theory of price determination. The ethical element implied in the natural prices was pushed into the background, and as a final result, entirely discarded. This development ended up in the 'Marginalist Revolution' of the 1870s.

Jean-Baptiste Say

In his *Traité d'économie politique*, the French economist Jean-Baptiste Say (1803), an enthusiastic follower of Adam Smith, attempted to rationalize the harmonious and self-regulating property of a market or exchange economy. His famous law, known as Say's Law, explains that 'supply creates its own demand' (C–M–C'). It implies that there cannot be an overproduction of all goods: any producer sells their product (C) against money (M), which in turn is spent on some other product (C') needed by the producer of commodity C. Thus, money is a pure medium of exchange, a veil covering the real transactions (C–C'): money is neutral. With the classical economists, David Ricardo in particular, Say's Law was given a macroeconomic interpretation: saving, consisting of profits, is always invested, thus heralding Smith's 'virtue of saving' – higher saving leads to larger investment levels, enhancing economic growth. Hoarding money is irrational, since it does not yield any revenue; given this, saving is always invested, directly or indirectly through banks.

Marginal analysis

As we move into the first half of the nineteenth century, marginal analysis was making its way through underground movements. The notion of marginal utility was developed, giving rise to Gossen's (1854/1983) two laws, namely the law of diminishing marginal utility and the law of equalization of the marginal utility–price ratios, implying utility maximization. Gossen's first law states that if an additional unit of some good, for example a slice of bread, is consumed, total utility increases, to an ever-smaller extent though; thus, additional (marginal) utility diminishes. According to Gossen's second law, total utility is maximized if, for each dollar spent, the consumer gets the same amount of additional (marginal) utility for the last unit of each good consumed. If for some good the consumer gets a higher amount of additional utility than for other goods, more of the former good will be consumed and, as a consequence, marginal utility provided by additional unity of this good diminishes; in contrast, consumption of other goods is reduced and the marginal utility provided by the last unit consumed increases. This adjustment process goes on until the ratio between marginal utility and price is the same for all goods. At this stage, total utility cannot be increased further, implying that total utility is maximized.

By contrast, the notions of marginal productivity and of marginal costs were still vaguely formulated around 1850. (The marginal productivity of a factor of production is the additional output added by an additional unit of that

factor, all other factors being held constant. Marginal costs are the supplementary costs incurred if an additional unit of output is produced.)

These 'marginalist' underground movements finally led to the 'Marginalist Revolution', which occurred broadly between 1870 and 1890 and provided the foundations of the currently still ruling neoclassical mainstream.

The neoclassical school

Among the great founders of the neoclassical school are Stanley Jevons, Carl Menger, Léon Walras and Alfred Marshall. Walras elaborated the general equilibrium model, where all markets are in equilibrium simultaneously, which gradually became the fundamental framework of neoclassical theory. General equilibrium is established by price changes with quantities given, hence already produced. Equilibrium prices have to be such as to equalize the quantities supplied and the quantities demanded for each good. Hence, Walras's basic model is a pure exchange model. Subsequently, production also becomes a matter of exchange. The available quantities of factors of production are combined in each enterprise in proportions such that each dollar spent on labour, land or capital yields the same additional output for a supplementary unit of these factors of production. This implies cost minimization or maximization of profits; which, in turn, goes along with the optimal allocation of resources, that is, the factors of production.

General equilibrium and partial equilibrium approaches

The Walrasian model is of immense complexity, as it describes the demand and supply relationships, based on rational behaviour, and equilibrium conditions across all markets, required to determine simultaneously all prices and quantities. Alfred Marshall then rendered Walras's model operable through the partial equilibrium approach, which deals with supply and demand conditions on specific markets. As a result, each market is considered in isolation in order to be able to set out very simply the way towards equilibrium between supply and demand. This is illustrated by the famous Marshallian 'cross',[1] which is reproduced over and over again in economics textbooks (Figure 3.1).

In Figure 3.1, the demand curve is downward-sloping owing to diminishing marginal utility. On the other hand, the supply curve is upward-sloping, as marginal costs increase when the quantity produced grows. The equilibrium price (p^*) is determined by the intersection of both curves.

Figure 3.1 The Marshallian cross

In a Walrasian framework, equilibrium is reached by price adjustments: if the price (such as p_1 in Figure 3.1) is higher than the equilibrium price, the quantity supplied (0e) will exceed the quantity demanded (0d) and the price will decline until equilibrium (given by price p^* and quantity x^*) is reached. While price adjustments are typical for neoclassical economics, Marshall's route to equilibrium, however, is based on quantity adjustments. For example, to the left of the equilibrium position (E in Figure 3.1), the 'demand price' (p_2 = ab in Figure 3.1), which consumers are willing to pay and producers would get for a certain quantity of goods supplied (0b), is above the 'supply price' (p_3 = bc) indicating the additional (marginal) costs of production of an extra unit. The difference between the two prices represents an additional profit for firms. Hence, entrepreneurs will produce more until the supply price and the demand price coincide at the point of intersection (E) of the two curves. At this point additional profits are zero and total profits are maximized (if the minimum of the average total costs is below the line p^*E), reflecting rational behaviour of producers.

Since with perfect competition there are a great number of firms in a given market, any firm cannot influence the market price (p^*) and can sell as much as it wants of its goods at this price. Hence, in neoclassical–Marshallian theory output must be determined by supply factors. Indeed, as long as the given market price exceeds the marginal costs incurred in producing an additional unit of the relevant good, the individual firm will expand its production until

marginal costs are equal to the (predetermined) market price (p^*). Hence, each firm maximizes the volume of its output, which is determined by scarce resources. This implication of neoclassical–Marshallian partial equilibrium analysis is of paramount importance, because output appears as being determined by supply factors and, as such, is supply-side economics. This is also characteristic of current mainstream economics. Marshall's *Principles of Economics*, published in 1890, thus became the 'bible' of economic theory. This work represents the basis for the great number of economics' textbooks written after World War II.

Within this Walrasian–Marshallian framework, all the great economic problems are solved on the basis of the principle of supply and demand with all markets being interlinked. As argued by Schumpeter (1954, p. 242):

> the all-pervading interdependence [of markets, hence of prices and quantities] is the fundamental fact, the analysis of which is the chief source of additions that the specifically scientific attitude has to make to the practical man's knowledge of economic phenomena; and that the most fundamental of all specifically scientific questions is the question whether analysis of that interdependence will yield relations sufficient to determine – if possible, uniquely – all the prices and quantities of products and productive services that constitute the economic system. [Walras's] system of equations, defining (static) equilibrium in a system of interdependent quantities, is the Magna Carta of economic theory.

This statement is of paramount importance and has greatly contributed to establishing the still ongoing supremacy of the neoclassical mainstream. Indeed, most neoclassical economists cannot imagine that one can theorize outside the Walrasian–Marshallian system.

The neoclassical synthesis

Neoclassical economics has dominated the theoretical scene from the 1870s onwards until now. Only in the great cyclical upswing, lasting broadly from 1950 to 1973, the year of the first oil price shock, had neoclassical economics made some concessions to Keynesian economics in the form of Paul Samuelson's neoclassical synthesis, which brings together Keynes's theory of employment determination through effective demand and Marshall's principle of supply and demand applied to the markets of factors of production and of final goods. A kind of mechanical Keynesianism in the shape of fiscal and monetary policies was to bring about full employment; given this, the goods markets would determine the prices of the final products and the associated

quantities, and the factor markets regulate distribution, that is, wages, profits and rents at the full employment level.

However, the neoclassical synthesis remained essentially neoclassical economics and, as a consequence, Keynes was to be integrated into neoclassical economics. In fact, in his *General Theory* Keynes claimed that effective demand erected a barrier to the left of the labour market equilibrium. As a result, full employment could not be reached. This becomes understandable if Figure 3.1 is considered to illustrate the labour market, where p represents the average wage rate and x the labour force. The effective demand barrier would be given by the line ab. In a Keynesian perspective, the equilibrium level of employment would be given by 0b and involuntary unemployment would equal bx^*. Entrepreneurs would not hire more than 0b workers, because the additional output produced by these workers, assisted by the existing capital stock or industrial plant, could not be sold, precisely because of limited effective demand.

However, from the 1970s onwards, neoclassical economists began to turn the tables. Positions to the left of the labour market equilibrium were explained by the average wage rate being above the equilibrium wage rate. Unemployment arose because wages (p_2 in Figure 3.1) were too high and had to be lowered in order to bring about a tendency towards full employment at (p^*, x^*). The principle of effective demand was eliminated and the way to fully restoring the reign of supply and demand was open.

Keynesianism indeed collapsed in the 1970s with the advent of inflation, owing to a sharp rise in the price of oil. In the neoclassical view, inflation made some of the basic core assumptions of Keynesianism unsustainable. Increasing the quantity of money, it was argued, did not result in lower interest rates, higher investment volumes and larger output and employment levels, but simply in higher price levels. Given this, from the mid-1970s onwards, monetarism, led by Milton Friedman, triumphed.

Monetarism and new classical economics

Monetarists are a group of economists who believe that competitive markets produce an inherent tendency towards full employment and that the quantity of money decided by the central bank determines only the price level. Monetarism was developed further in the Walrasian new classical model, where prices are entirely flexible. As a result, there is a continuous market clearing, that is, the quantity supplied always equals the quantity demanded on each market. In regard to Figure 3.1, we are always in a (p^*, x^*) equilibrium

situation in all markets. However, the behaviour of producers, consumers and states (economic agents) has consequences for the future. Economic agents attempt to come to grips with future developments through taking into account all available information: for example, the best available economic theory (the general equilibrium model), the possible evolution of the oil price, future fiscal policies and the evolution of public debt.

However, economic activity is influenced from time to time by unforeseeable external shocks, which may occur either on the demand side (important shifts in demand for some goods) or on the supply side (profound technological changes: the digital revolution, for example). Supply and demand shocks lead to cyclical movements in economic activity; that is, movements in the scale of output and employment.

The new classical economists postulate that in the course of business cycles the economy is always in equilibrium too: at market clearing prices, the quantities supplied always match the quantities demanded. Most importantly, there is no involuntary unemployment (which is evident if Figure 3.1 is considered as representing the labour market). All those out of work are voluntarily unemployed, that is, they do not want to accept a job at the prevailing wage rate. Consequently, new classical economics, led by Robert Lucas, is entirely supply-oriented. The actually existing quantities of production factors govern economic activity; that is, output and employment. This equilibrium vision of the economy is strictly anti-Keynesian. Indeed, with Keynes, effective demand in money terms by consumers, producers and the state governs economic activity. It increases as the distribution of income becomes more equal, which in turn enhances the spending power of the population. Hence, Keynes and Lucas are, in fact, the great antagonists in modern economic theory.

New Keynesian economics

While new classical economics is equilibrium economics, new Keynesian economics is disequilibrium economics. The new Keynesians assume that an equilibrium position does exist (p^*, x^* in Figure 3.1). A disequilibrium (such as p_2 and the quantity 0b in Figure 3.1) may occur and persist, because there is no Walrasian auctioneer to establish the equilibrium price. Indeed, disequilibrium situations may persist for various reasons. For example, the volume of production (0b in Figure 3.1) cannot be increased owing to a persistent shortage of some factors of production or because of a supply cartel, which restricts output. Persistent disequilibria may also exist because some institutions obstruct the functioning of perfect markets as pictured by the

demand and supply curves in Figure 3.1. If Figure 3.1 represented the labour market, trade unions would try to shift the labour supply curve upwards and to the left; entrepreneurial associations would attempt to shift the labour demand curve to the left and downwards. A new equilibrium would eventually come into being far to the left of the original equilibrium (E), thereby implying an employment level far below the full employment level ($0x^*$). All in all, imperfect competition, trade unions, cartels, monopolies and oligopolies may bring about persistent unemployment.

The history of political economy

Considering the history of political economy means leaving the frictionless machine of economics to enter an altogether different theoretical world. What is put to the fore is not the market but the social and circular process of production. The social product – gross domestic product (GDP) – results from a common effort within all industries. Firms can produce only because they receive goods from other firms. For example, the shoe factory receives leather and specific machines; on the other hand, firms producing primary goods (such as iron, steel, electricity and machine tools) deliver their commodities to firms producing intermediate goods and final goods. Hence, the social process of production is a complex system of deliveries and receipts, which is represented by the so-called Leontief quantity system. Given the social nature of production, all the great problems of economic theory – value and price, distribution, employment, money and finance – are social and macroeconomic issues. We will expand on this later.

Political economy also rests on a specific vision of society. The heart of the economy is the social process of production and monetary and financial institutions (the central bank and the banking system). The economy is a monetary production economy that forms the material basis of a society and produces the social surplus, that is, the social product (or GDP) minus the socially necessary consumption by producers – namely, workers.

The social surplus is used up to maintain and to expand the production system through gross investments and, most importantly, to build up political, social, legal and cultural institutions as well as a comprehensive education system. These institutions form the institutional superstructure, and together with economic institutions (enterprises in the social process of production and the banking system) they form a complementary set of institutions. Given the complementarity of institutions, the material basis of a society and its institutional superstructure form a system or a structured entity, which may be called the socioeconomic system or society for short.

In political economy, society is primary and the starting point of analysis. Given this, political economy is essentially about the functioning of the socioeconomic system, not about the rational behaviour of individuals, which is supposed to be coordinated in a socially meaningful sense by competitive markets.

François Quesnay

The political economy strand starts with the French surgeon François Quesnay (1694–1774), whose *tableau économique fondamental* (fundamental economic table) (Meek, 1962, p. 275), first published in 1758, represents for the first time in the history of economic theories the social and circular process of production with flows of goods moving between industry and agriculture, and flows of money moving in the opposite direction to buy the goods. Hence, goods are always exchanged against money, never against other goods. Agriculture delivers necessary consumption goods and raw materials to industry (handicrafts and manufactures), which in turn provides agriculture with industrial consumption goods (cloth and shoes, for example) and various tools (investment goods).

In Quesnay's model, prices are determined within the process of production. These prices of production are known before goods arrive on the market, that is, before they enter the process of circulation, as opposed to being determined by supply and demand once they arrive on the markets. All prices are made up of the costs of production and of a profit or rent element.

The expenditures of the state are of the greatest importance for Quesnay. As a physician he considers the state as the heart of the socioeconomic system: through its spending, the state injects money (blood) into the socioeconomic system (body); money is reproduced by the agricultural sector (the stomach) and returned to the heart (the state) in the form of land rent, thereby providing revenues to the state. This emerges from Quesnay's *tableau économique* (Figure 3.2), presented here in an elaborated form. In fact, Quesnay's original figure contains the three central columns (III, IV and V) only (Meek, 1962, p. 275) and, as such, is not easy to understand. To render Quesnay's *tableau* intelligible, columns I, II and VI have been added in Figure 3.2.

Quesnay's *tableau* represents the French economy around 1750 in its sound and natural state. To be able to describe its functioning, a glance at the structure of the French society is required. In pre-revolutionary, eighteenth-century France, the landowners, the nobility, the King and the upper strata of the bourgeoisie also make up the state, which, in addition, comprises the army

Figure 3.2 Quesnay's *tableau économique*

and the state administration, including the legal system. The state revenue (land rents and taxes) amounts to £2000 and appears at the top of column IV. Column III stands for agriculture. Here the tenant farmers dispose of advances in the form of wheat (necessary consumption goods) with a monetary value of £2000. These advances are required to feed the agricultural labour force for the coming production period (a year). The manufacturing sector will use its advances – £1000 in money form (top of column V) – to buy raw materials from the agricultural sector (column II).

The economy, that is, the processes of social production and circulation, is set into motion through the spending of the revenues of the state or the landlords (£2000). Half of the state expenditures (£1000) is spent on agricultural goods (column III), half on manufactured goods (£1000, column V). This demand initiates production with a money value and corresponding income amounting to £1000 in both sectors. Agricultural goods (necessary consumption goods) worth £1000 and manufactured goods worth £1000 now flow to the state to maintain the King and the nobility, government and administration as well as the army. Subsequently, tenant farmers

spend half of the newly produced income (£500) on manufactured goods (consumption and (replacement) investment goods, column V); the other half (£500) is put aside in money terms (column I). Similarly, the manufacturing sector buys agricultural goods (necessary consumption goods) from agriculture (£500 in column III), putting aside £500 in view of paying for the raw materials of the next year (column VI). The process goes on as suggested in Figure 3.2.

The surplus column IV remains to be explained. This requires a careful consideration of the figures in column III (£1000, £500, and so on), which have three different meanings. In the first place, these figures stand for the expenditures in money terms of the state (£1000) and of the manufacturing sector (£500) on agricultural (necessary consumption) goods. Second, these amounts represent the money value of inputs in the (productive) agricultural sector. Indeed, the tenant farmers use the sums of money in column III they have received from the landlords (the state) and the manufacturing sector to pay wages to the agricultural labour force and to buy seeds. The agricultural workers and the tenant farmers spend these amounts of money on the advances of the agricultural sector (wheat worth £2000 at the top of column III). Now, production in the agricultural sector may start. An input of £1000 (labour and seeds) leads to an output of £2000, an input of £500 yields an output of £1000, and so on. Hence, in the third meaning, the figures in column III represent outputs, replacing the inputs incurred. Moreover, owing to the forces of nature, a surplus in terms of wheat amounting to £1000, £500, and so on, arises (£2000 in column IV), which provides the advances of the agricultural sector for the next year (top of column III, dotted line a). The surplus in money terms arises in column I (£2000) and consists of half of the sales receipts from the state (£500) and from the sales to the manufacturing sector (£1500): £500 for food sales (columns III and V from the third line onwards) and £1000 for the sale of raw materials. Indeed, the manufacturing sector spends its advances of £1000 in money terms (top of column V) to buy raw materials from the agricultural sector (column II, dotted line b). The surplus in money terms (column I) is used to pay the rents to the landlords (the state); these revenues provide the advances of the state to be spent next year (£2000 at the top of column IV).

Quesnay rightly argues that a surplus can arise in agriculture only, where the goods necessary in the social process of production are produced. In the (sterile) manufacturing sector no surplus can arise. Here output is worth £2000 (column V) and the inputs are delivered by agriculture: raw materials worth £1000 (columns II and V) and food, also worth £1000 (columns III and V). In the manufacturing sector competition is so intense as to reduce

profits to zero; given this, the value of inputs (£2000) equals the value of output, also £2000 (column V).

Quesnay has been criticized for postulating that land (agriculture) is productive, not labour. This is a misunderstanding. In fact, in the social process of production, land and labour are complementary. Marx said that human beings (labour) interact with nature (land) by means of tools and machines. Quesnay looked at production from the land perspective, Ricardo from the labour perspective. The different points of view taken by Quesnay and Ricardo arise from the different problems both were dealing with. Quesnay considered the issues of the scale of economic activity – the level of output (and employment) – and distribution; Ricardo dealt with the issues of value and distribution.

Output and income are thus produced in both the agricultural and manufacturing sectors, enabling both sectors to buy goods from each other. This interaction between both sectors brings the social nature of production, circulation and consumption to the open. This mutual spending of income represents a cumulative process, resulting in the production of agricultural goods worth £5000 in money terms (columns II, III and IV in Figure 3.2) and of manufactures amounting to £2000 (column V). The total output of £7000 is the maximum output, implying full employment. Hence, in a natural and healthy state the appropriate distribution and spending proportions result in the maximum scale of economic activity. Based on the healthy state of the economy, surgeon Quesnay is able to deal with states of economic illness. Most importantly, he worries about landowners not spending the entire amount of their rent (£2000), but engaging in unproductive speculative activities or simply hoarding. Obviously, if the landowners spend less than £2000, economic activity and employment would be reduced, thereby inducing system-caused involuntary unemployment. In 1930, Keynes said that crises arise because money is flowing from the real sector, where production takes place, to the financial sector, in which no income-generating transactions occur. Both Quesnay and Keynes are thereby brothers in spirit.

To be sure, there is no self-regulation of the economic system in Quesnay's *tableau économique*. For Quesnay, the level of economic activity and employment is a political issue, associated with the level of government expenditures and the distribution of national income. Socially appropriate proportions between agricultural rents or the social surplus, the revenues of the farmers, and the wages of the agricultural labour force as well as the industrial wage bill have to be established. For example, the wages in agriculture must be such as to use up the advances in agriculture (£2000). Quesnay explicitly

insists on the necessity of paying fair and good wages to agricultural labourers. Not using up the advances would reduce production and bring about a crisis situation.

Classical–Keynesian political economy

The classical–Keynesian supermultiplier to be dealt with at the end of this section, the Leontief quantity system (deliveries and receipts of goods) and Sraffa's prices of production (Sraffa, 1960, p. 11) are straightforward developments of Quesnay's *tableau*. Hence, classical–Keynesian political economy (see Bortis, 1997, 2003, 2013, 2015) is the elaborated form of Quesnay's *tableau économique* in modern dress. This is perhaps the most important fact in the history of political economy. François Quesnay indeed appears as the founder of political economy.

David Ricardo

Adam Smith, who pushed the social process of production and the state into the background, radically changed the picture, and the self-regulating market now eclipses production. With David Ricardo, however, the pendulum of economic theory swung back to political economy: the functioning of the socioeconomic system moves to the fore at the expense of the behaviour of individuals. The social process of production moves to the fore again, with exchange or markets – so important for Adam Smith – being pushed into the background. In contradistinction to Quesnay, however, Ricardo considered the labour aspect of social production. The value of commodities is now governed by direct and indirect labour time; this is what is called the labour theory of value. In other words, what gives value to goods is the amount of labour (or labour time) used in their production. Labour can then be divided into direct and indirect labour. Direct labour is used to produce final products; indirect labour is used for producing the various means of production, namely, used up fixed capital, intermediate and primary goods. Hence, labour values represent the social effort made to produce any commodity.

However, the main problem Ricardo considered is not value but distribution. According to him:

> The produce of the earth – all that is derived from its surface by the united application of labour, machinery, and capital, is divided among the three classes of the community; namely, the proprietor of the land, the owner of the stock or capital necessary for its cultivation, and the labourers by whose industry it

> is cultivated . . . To determine the laws which regulate this distribution, is the principal problem in Political Economy. (Ricardo, 1821/1951, p. 5)

These laws consist of two principles that operate in agriculture, where necessaries, represented by corn, are produced.

First, there is the marginal principle, which determines agricultural rents. The least fertile land just cultivated gets no rent. This is a simplifying assumption required in order to establish the labour value principle. Indeed, with the rent on the least fertile (marginal) land eliminated, the cost of producing corn is determined by direct and indirect labour costs only. This highlights the ingenious method used by Ricardo. He was in fact dealing only with essentials or fundamental causal forces, namely, principles that are absolutely necessary to explain a phenomenon; the value of produced goods in the present instance. Hence, the title of his main work: *On the Principles of Political Economy and Taxation* (Ricardo, 1821/1951). In models dealing with fundamentals or principles:

> [only] what is considered to be essential or constitutive to a phenomenon is included in the model, which is a picture, in fact a *reconstruction* or *recreation* of what . . . *constitutes* a phenomenon (for example, prices, quantities and employment levels in political economy). This recreation is performed by reason interacting with intuition and is analogous . . . to the representation of essential information for the user of the underground through a map. (Bortis, 2003, p. 413)

Joan Robinson (1962, p. 33) also compared a theory with a map, adding that a realistic map where the entire reality is represented would be totally useless. Ricardo is very important for political economy, because he was the first to build a logically consistent theoretical model, made up of essentials, in order to be able to explain the immensely complex real world by simple causal models which, in turn, provide the basis for policy prescriptions.

Hence, Ricardo eliminates rent, crucial for Quesnay, from value formation. This can be generalized: labour values govern prices where the conditions of production are most difficult. This gives rise to various types of rent, occurring when the conditions of production are more favourable.

Now, with rent determined through the marginal principle, another distributional principle, namely the surplus principle, comes in to determine wages and profits in agriculture in real terms, that is, in terms of corn. Ricardo first postulated that the real wage rate is not governed by market forces – that is, by supply and demand – but by social and cultural factors:

> The power of the labourer to support himself, and the family which may be necessary to keep up the number of labourers, does not depend on the quantity of money which he may receive for wages, but on the quantity of food, necessaries, and conveniences become essential to him from habit, which that money will purchase. The natural price of labour, therefore, depends on the price of the food, necessaries, and conveniences required for the support of the labourer and his family. (Ricardo, 1821/1951, p. 93)

In this view, profits are merely the surplus over wages of the agricultural output, net of rent. Interestingly, in the view of David Ricardo profits were justified only if they were invested.

In his corn model (Ricardo, 1815) the rate of profits is given by the ratio of corn profits to the agricultural wages sum (circulating capital), also in terms of corn, which stands for necessary consumption goods. This implies two things. First, the agricultural labour force produces not only corn, but also the tools required in production, fixed capital (machinery) being absent. Second, distribution must necessarily be regulated in the sectors producing basic or socially necessary goods, because the natural wage rate, and hence the wages sum (circulating capital), are in terms of corn. Given this, the (natural) agricultural profit rate determines the rate of profits in the sectors producing luxuries, silk for example.

Ricardo's corn model was criticized on the grounds that, in an industrial economy, corn was not the only good, because there were different consumption and capital goods. Moreover, fixed capital was gaining in importance. A new measure had to be found for the wage goods and for calculating the rate of profits. In his *Principles*, Ricardo proposes measuring the value of all produced goods by direct and indirect labour, which in a way replaces corn. This greatly widens the scope of Ricardo's theory of income distribution, essentially based on the surplus principle. Potentially, distribution becomes positively a matter of social power and normatively an issue of distributive justice.

Reactions against Ricardo and Marx

Given the 'dangerous' and political implications of his theory of value and distribution (recall the discussion in Chapter 1), the reaction against Ricardo set in almost immediately after his death (Dobb, 1973, p. 96) in the form of Smithianism, that is, still vague demand and supply theory. While economics – the precursor of the Marginalist School – grew stronger underground, suddenly in 1867 a powerful piece of political economy was published; that is, the first volume of Karl Marx's *Das Kapital* (Marx, 1867–94/1973–74),

which placed Ricardo's *Principles* in a very wide and profound framework of politics and history. In the mind of many of the time, the two books were considered to be cut from the same theoretical cloth.

This ideological factor was certainly a crucial element in setting off the Marginalist Revolution. In fact, Ricardo and Marx were unable to provide a satisfactory solution to the so-called 'transformation problem': that is, the problem of transforming labour values into prices of production. The difficulty arose, first, because of the growing importance of fixed capital goods, embodying past labour and having a durability of several years; and second, because the rate of profits had in principle to be uniform in all sectors of production to enable the classical mechanism of competition to work properly: capital would always flow to sectors where the realized rate of profits durably exceeded the normal rate of profits, and vice versa. Now, the problem arises if the conditions of production, given by the ratio of the (past) labour embodied in fixed capital to (present) labour embodied in prime costs, is unequal in the various sectors of production (textiles, cars, and so on). Given the uniform rate of profits, the prices of production will exceed the total labour – past and present – embodied in the production of some good if its production is capital-intensive, and vice versa. Hence, the prices of production are not proportional to labour values (wage costs) but deviate from them in an unpredictable way. Ricardo and Marx were unable to explain how labour values were transformed into prices of production (this problem was solved by Piero Sraffa in his 1960 book *Production of Commodities by Means of Commodities*; see the portrait of Sraffa at the end of Chapter 2). Neoclassical economists considered that classical value theory had moved into a blind alley, and as a consequence decided to shift the explanation of value from the social process of production to the market, governed by the marginal principle embodied in the law of supply and demand. This is the essence of the Marginalist Revolution.

Marshall's *Principles of Economics* (first published in 1890) thus became the 'bible' of economics and the basis of the abundant neoclassical textbooks literature. Subsequently, classical political economy along the lines of François Quesnay and David Ricardo had 'been submerged and forgotten since the advent of the 'marginal' method' (Sraffa, 1960, p. v).

A classical-Keynesian counter-revolution

In the interwar years 1926–36, Piero Sraffa and John Maynard Keynes produced a theoretical twin revolution, in fact a classical–Keynesian counter-revolution against the utterly dominating neoclassical school. In the mid-

1920s Sraffa attacked Marshall's supply curve and argued that marginal costs were not rising, but remained constant when output varied, because capacity utilization increases and money wages are fixed by contracts. Most importantly, however, money wages do not rise when additional workers are hired, because of ever-existing unemployment, which exerts a continuous downward pressure on wages. With marginal or prime costs given, firms may price their goods based on what is called a mark-up approach: prices are determined by imposing a mark-up over costs of production. To have a benchmark price, the mark-up is imposed on prime costs at normal capacity utilization in order to cover fixed costs and to bring about a normal rate of profits. Market prices and realized profits will, as a rule, deviate from these (normal) benchmark prices, and actual capacity utilization will differ from normal capacity utilization. If realized profit rates are persistently below the normal profit rate, firms tend to cut back capacities, and vice versa. This is the classical view of competition: the prices of production are determined in the production process and demand determines the quantities produced.

In his 1960 book, Sraffa shows how the prices of production are determined and distribution regulated in principle within the social and circular process of production (see particularly p. 11). With the normal prices determined, demand determines the quantities that can be sold. Given this, Sraffa has provided the microeconomic foundations for Keynes's theory of aggregate output and employment as governed by effective demand.

Indeed, in *The General Theory of Employment, Interest and Money*, John Maynard Keynes (1936), the second great protagonist of the twin revolution of the interwar years, for the first time convincingly challenges Say's Law, stating that economic activity is governed by supply factors, that is, scarce resources in the form of labour, land and capital. The Keynesian revolution represents the monetary way to the principle of effective demand (Garegnani, 1983), exhibited by the multiplier relation, explaining how both output and employment are determined in principle:

$$Q = (1/s)\, I = [1/(1 - c)]\, I \tag{3.1}$$

The economy is set into motion through investment expenditures (I), which subsequently bring about a cumulative process of demand and production of consumption goods. The equilibrium level of output determined by the multiplier $(1/s) = [1/(1 - c)]$ may imply an employment level well below the full employment level ($Q < Q_f$). System-caused involuntary unemployment comes into being. Contrary to Say's Law, general overproduction – that is, overproduction of all goods – is therefore possible.

The principle of effective demand could be established because saving and consumption mainly depend on current income, and the rate of interest has but a minor and unpredictable effect on them. The rate of interest has the new task of bringing into line the exogenously given amount of money with the demand for money. The supply of money is determined by the central bank. The demand for money is twofold: first, for transaction purposes in the real economy where newly produced goods are always exchanged against money; and second, for speculative purposes: money is held because of uncertainty about the future course of real and financial assets. For example, when the Dow Jones reaches very high levels, people tend to remain liquid, refraining from buying shares, and vice versa.

Outline of the classical–Keynesian system

Above we have suggested that Keynes has been absorbed by neoclassical theory in the form of new Keynesianism. Where then do we stand today? Is there an alternative to this neoclassical–Walrasian exchange model? The answer lies precisely in elaborating on and then merging together the insights of Keynes with those of Sraffa as well as Leontief to form a classical–Keynesian monetary production framework (see Chapter 4 for a full discussion).

Such an alternative is captured most appropriately by Marx's scheme of production and circulation of capital (Marx, 1867–94/1973–74, Vol. II, p. 31):

$$M–C \ldots P \ldots C'–M' \tag{3.2}$$

where M represents money and finance (financial sector), C the means of production, P the social process of production, C' final output (social product), and M' money (effective demand). Keynes, implicitly relying on Marx, explicitly aimed at working out a monetary theory of production to be able to explain the deep crisis of the 1930s. His efforts resulted in *The General Theory*, which is mainly about the sequence $C'–M'$ (effective demand governing output and employment). In earlier works, Keynes dealt with $M–C$ (money and banking); Wassily Leontief was concerned with the quantity flows within the social and circular process of production (P); and, finally Piero Sraffa, in close touch with David Ricardo and Karl Marx, considered the principles of price formation and the regulation of distribution within this process. Subsequently, Luigi Pasinetti has decisively contributed to preparing a synthesis of these authors and their followers, thus clearing the way towards a classical–Keynesian system of political economy (Bortis, 2012).

The basic classical–Keynesian model must be made up of the primary principles governing the functioning of modern monetary production economies. Two principles are of classical origin: the labour value principle summarizes the essential features of the immensely complex social process of production to provide the essence of the prices of production, which are the fundamental prices in a monetary production economy (Bortis, 2003, pp. 433–45); the surplus principle of distribution implies that the distribution of income is positively a problem of social power and normatively of distributive justice situated at the heart of social ethics (Bortis, 1997, pp. 158–75). Keynes provided a third principle, namely the principle of effective demand, related to determining the scale of economic activity (Bortis, 2003, pp. 460–67). These three principles imply that money, intimately associated with the financial sector, plays a fundamental role. Indeed, the processes of production and circulation could not go on without money, since production takes time, outlays and receipts are not synchronized, and goods are never exchanged against other goods (as is the case in a neoclassical–Walrasian framework) but always against money, which also acts as a store of value and as such is intimately connected to the financial sector.

It is of the utmost importance to bring together these principles in a coherent theoretical framework that may be set into opposition to the neoclassical–Walrasian framework. Indeed, as emerges from Keynes's economic and philosophical work, to act on the basis of principles is the most appropriate way to act rationally in a complex and rapidly evolving real world of which we have imperfect and probable knowledge only, and where uncertainty about the future always prevails. The great problem is to uncover the most plausible principles on which to base our actions. To make economic analysis fit for purpose requires working out a fundamental classical–Keynesian system of pure theory to bring into the open how monetary production economies essentially function and to compare this theoretical system with the basic neoclassical–Walrasian model.

In fact, it is not sufficient to simply establish the system of classical–Keynesian political economy: the neoclassical theory must still be proven wrong. The neoclassical real theory (that is, the law of supply and demand) has to be attacked at its foundations. This attack was carried out in the course of the Cambridge capital-theoretic debate (Harcourt, 1972). Since this highly important, but immensely complex debate cannot be presented here, only the bare essentials of the argument are set out in what follows.

The capital-theoretic debate

The starting point is the measurement in physical terms of the factors of production (labour, land and capital) as is required by the concept of marginal productivity of a factor of production, essential in neoclassical economics. For example, the marginal product of labour is given by the additional output produced by employing an additional worker for a given unit of time, say a week. Now, the natural factors of production (land and labour) can be measured physically, in acres and labour-hours. However, it is quite evident that it is impossible to measure real capital (machines and factory buildings) physically; in tonnes of steel, for example. Hence, capital must be measured in money terms. But since capital is a produced factor of production, this requires knowing Sraffa's prices of production. This means, in turn, that the money wage rates and the target profit rate must be known (see also the price equation 3.3, below). Neoclassical economics is now faced with a contradiction: on the one hand, the rate of profit and money wages, hence the normal prices of capital goods, must be known in order to be able to measure the value of capital in money terms; on the other hand, the rate of profits must be an unknown to be determined on the markets for new capital goods. Here saving, increasing as the rate of interest rises, represents the supply of new capital; and the volume of investment, increasing as the rate of interest decreases, stands for the demand for new capital. This is the well-behaved downward-sloping demand curve for new capital goods: investment increases as the rate of interest decreases.

Now, the basic result of the capital-theoretic debate is that, in principle, no well-behaved associations between the volume of investment and the rate of interest need necessarily exist. This implies that the principle of supply and demand is not compatible with the principle of effective demand, since in principle the volume of investment cannot necessarily adjust to full-employment saving in the long run, with profit rates being equal in all sectors of production (Garegnani, 1983). However, Keynes's *General Theory* is based on the fact that, in macroeconomic equilibrium, saving always equals investment. This leads to the Keynesian multiplier (equation 3.1) and to the classical–Keynesian supermultiplier (equation 3.4, below). The way to effective demand, Garegnani's real way, is thereby definitely cleared and classical–Keynesian political economy may be confidently established.

In the remainder of this chapter we will touch upon two fundamental issues: the formation of prices in relation to income distribution, and the determination of the level of output and employment. However, before we start, some methodological issues are to be dealt with.

Preliminary remarks on method: principles and pure theory, and applied theory

The distinction between principles and pure theory on the one hand, and applied theory on the other, is of the greatest importance in the history of economic theories. Pure theory is based on principles. For example, the price equation (3.3) is the pure theory of fundamental price based upon the labour value principle; this equation represents what is essential about prices; consequently, all the elements that would modify the fundamental labour value price are abstracted from: the conditions of production (the ratio between fixed capital and labour force), the influence of business cycles on prices and price movements due to the (short-term) vagaries of the market. Similarly, the supermultiplier relation (3.4) represents the pure theory of the institutionally governed output (and employment) trend, based upon the principle of effective demand; again, the influence of business cycles and market movements on output and employment is abstracted from.

Principles and pure theories are intellectual constructs or products of the mind; they recreate or reconstruct by means of reason what is considered to be constitutive or essential of a phenomenon; for example, prices, quantities, and output and employment levels.

Applied theory, however, deals with objectively given phenomena, existing in space and historical time. For instance, the attempt to explain the evolution of output and employment in the United States after World War II until the present would represent a piece of applied theory. Evidently, pure theory provides the tools, the concepts, put to use in applied theory.

This chapter uniquely deals with the principles and pure theory, simply because it is pure theory that is fundamental and, therefore, constitutes the main object of the history of economic theories. This methodological device carries several implications.

In the first place, the very simple relations presented in this chapter – (3.1), (3.3), (3.4) and (3.5) – should be considered a kind of shorthand writing. This allows to set forth very complex macroeconomic states of affairs with far-reaching implications in a simple, concise and precise way. It would be very difficult to describe these complex situations in plain words with sufficient clarity. Second, the equations of this chapter are very simple macroeconomic causal relations: the independent variables on the right-hand side determine the dependent variable on the left-hand side. As such, the classical–Keynesian causal models stand in striking contrast to the neoclassical equilibrium models where the equilibrium is the result of two opposing

forces, supply and demand, on various markets; moreover, there is also a striking contrast between the very simple macroeconomic classical–Keynesian causal models and the immensely complex, microfounded, general equilibrium model, which is the basic model of neoclassical economics. Third, as suggested above, the subsequent equations represent pure theory, picturing fundamentals or essentials, which implies that all the independent and dependent variables are of an abstract nature, devoid of physical properties; in philosophical language, pure theory and the underlying principles are of a metaphysical nature.

Given this, the price equation (3.3) implies that labour values represent the essence of the prices of production, which depend on the conditions of production, that is, differing relations between fixed capital and the labour force; indeed, the conditions of production do not produce values; as Ricardo said, they only modify values and can therefore be left aside. The same holds for the influence of business cycle movements and the vagaries of the markets, both of which are abstracted from. Fourth, in pure theory, all the independent variables are independent of each other. This also holds for the trend and business-cycle variables picturing deviations from the trend (relations 3.4 and 3.5 and Figure 3.3). Moreover, time in Figure 3.3 is not historical time, but logical time (Joan Robinson). Hence pure theory is universal and ahistorical, that is, it holds everywhere and at any time. Applied theory, however, considers historical realizations of the variables and parameters contained in

Figure 3.3 The supermultiplier relation

pure theory. Here the variables and parameters are interdependent and the process of causation goes on in historical time. Consequently, empirical and historical analyses, as a rule, become very complex. However, pure theory greatly contributes to understanding what is going on in the real world; in fact, pure theory illuminates empirical and historical facts and, therefore, enhances our understanding of real-world phenomena, such as price formation within firms, employment levels in various countries and distributional outcomes worldwide. Fifth, pure theory also represents the basis for economic policies, for instance the necessity for incomes policies to enhance the purchasing power of the population, which, in turn, increases the level of output and employment (in relation 3.4, this would be reflected in a higher trend consumption-income ratio c^* and in a larger trend output Q^*). In this context, Keynes argued that in a complex and rapidly evolving real world, the most sensible thing policy-makers can do is to rely on principles. Both Ricardo and Keynes were masters in this domain, confirming thus a well-known saying: 'Nothing is more practical than good (pure) theory'.

Hence, to consider principles embodied in pure theory is of the utmost importance for economic theory and policy, as David Ricardo was the first to clearly perceive.

The formation of prices

The formation of prices and the regulation of distribution are of immense complexity within the nature aspect of production, considered by Piero Sraffa (Pasinetti, 1977, pp. 71–151). Here the prices of production depend on all the production coefficients making up a technique of production and on income distribution, that is, the money wage rate and the rate of profits. However, moving on to the labour perspective of production leads to a very simple model of value and distribution that can be used for macroeconomic purposes (Bortis, 2003, pp. 436–45). Indeed, each price (p^*) is now given by the product of the money wage rate (w_n), the quantity of direct and indirect labour (N) per unit of output ($n = N/Q$) and the mark-up ($k(r^*)$), which has to ensure a target rate of profits on fixed capital (r^*), which is firmly anchored in the institutional system, a point made by Pierangelo Garegnani time and again. This sectorial price equation leads on to a macroeconomic price equation, based on the labour value principle:

$$p^* = w_n nk = w_n\,(1/A)\,k(r^*) \tag{3.3}$$

Equation (3.3) captures what is essential about the prices of a monetary production economy (summarized by the macroeconomic price of production);

the influence of business cycles and of market vagaries on prices are abstracted from. Overall labour productivity, A, is the inverse of the macroeconomic labour coefficient, n, with $A = Q/N$ and $n = N/Q$, where N is the (direct and indirect) productive labour force active in the 'profit sector' for specified work hours. The social product, Q, is measured in terms of a bundle of socially necessary consumption goods, of which p^* is the normal price or price of production in terms of money. The coefficient n tells us that in a monetary production economy there are labour costs only, since ultimately labour in the expression $(w_n n)$ represents the unit cost in terms of money. As a result, the amount of direct and indirect labour time expended in producing a unit of the social product governs the value of this unit in a Ricardian vein. In a Marxian perspective, the price expresses the value of a product in terms of money. And prices are determined as soon as the money wage rate (w_n) and the mark-up $(k(r^*))$ are fixed. This immediately emerges from equation (3.3).

Hence the macroeconomic price of production (equation 3.3) increases as the money wage rate and the rate of profits increase, and declines as labour productivity increases (or the labour coefficient diminishes). However, equation (3.3) implies that there is an inverse relationship between the real wage (w_n/p) and the rate of profits (r^*). And, very importantly, profits are produced by labour, since in a classical–Keynesian perspective labour is, in a Ricardian vein, the only factor of production. This has crucially important implications for the interpretation of empirical facts. For example, the increasing inequalities are not due to a higher productivity of capital as some economists would argue. For the classical–Keynesian political economist the main cause of growing inequality is given by the existence of involuntary unemployment, which exerts a permanent pressure on real wages.

The money wage rate w_n and the mark-up $k(r^*)$ regulate income distribution, that is, the social surplus (surplus wages, land rents, labour rents owing to social power or to privileges, and, last but not least, profits) over socially necessary wages. Subsequently, the distribution of socially necessary wages and, above all, the distribution of the social surplus among the various social classes becomes a most fascinating problem of political economy, sociology and politics. In general, with given money wages, a larger $k(r^*)$ means that the surplus over socially necessary wages increases, that is, the distribution of income gets more unequal. This has important implications for the classical–Keynesian theory of output and employment to which we now turn.

The determination of long-period or trend output and employment: the supermultiplier

The classical–Keynesian long-period theory of output and employment is given by the supermultiplier relation. The star attached to all independent variables and parameters indicates that these are trend values governed by slowly changing factors, that is, technology and institutions:

$$Q^* = \frac{G^* + X^*}{(1 - c^*) + b^* - a^*} \tag{3.4}$$

This relation governs the position of the normal, trend or 'equilibrium' output, Q^*, which may be located well below the full employment trend Q_f, as Figure 3.3 shows.

The supermultiplier equation (3.4) represents the pure theory of output (and employment), based upon the principle of effective demand. As such, this equation provides the basis for applied theory – for example, the study of the evolution of output and employment in particular countries – and provides the foundations for economic policies, most importantly employment and distribution policies.

The distance between Q_f and Q^* in Figure 3.3 indicates permanent (long-period) involuntary unemployment. The existence of system-governed permanent involuntary unemployment is the most important feature of classical–Keynesian political economy. It is important to note that the supermultiplier determines the level or the volume of employment only. Who is employed or unemployed is more or less uncertain. And who gets jobs, above all the well-paid and very well-paid jobs, is a fascinating question of economic sociology.

In the supermultiplier relation (3.4), Q^* is real-trend GDP (equal to domestic income) measured in terms of a bundle of necessities having the price p^* (see equation 3.3). The right-hand side of the supermultiplier equation (3.4) represents effective demand, which governs output and employment. G^* stands for normal government expenditures, resting on parliamentary or government decisions. In highly industrialized countries trend exports X^* are governed, among other factors, by the quality of the education system and by research and development expenditures; c^* represents the trend propensity to consume out of income (C/Q) determined by evolving consumption habits.

Consequently, $(1 - c^*)$ equals the sum of the saving–income ratio $(s = S/Q)$ and the tax rate $(t = T/Q)$; S and T are the amounts saved and paid on taxes

respectively. The normal import coefficient (M/Q) is b^* (where M is the volume of imports) and indicates the technological and cultural dependence from the rest of the world. The gross investment–income ratio (I/Q) = a^* is the vehicle through which technical progress is diffused in the form of new and replacement investments (I). The trend rate of growth of effective demand and of GDP (Q^*) is given by the rate of growth of the autonomous variables ($G^* + X^*$). In light of its definition ($Q = AN$), potential (full employment) output (Q_f) grows at the natural rate of growth, given by the growth rate of labour productivity (A) – due to technical progress – and the labour force (N). Hence, if the growth rate of effective demand is lower than the natural rate of growth, involuntary unemployment increases. The condition $S = I$ represents the macroeconomic equilibrium condition, which, in a Keynesian vein, must always hold.

The institutionally governed autonomous expenditures – trend government expenditures (G^*) and trend exports (X^*) – set the economy into motion, resulting in a cumulative process of production of both consumption and investment goods, and simultaneously of income, part of which is spent abroad, giving rise to imports. Hence, the trend volumes of consumption (C^*), of imports (M^*) and, most importantly, of investment (I^*) are all derived magnitudes and determined by long-period effective demand. However, individual investment projects are subject to uncertainty about the future: which investment projects are successful, which ones fail? The fact that long-period (trend) investment $I^* = a^* Q^*$ is determined by trend output, governed, in turn, by long-period effective demand – the right-hand side of equation (3.4) – is highly important: trend investment I^* is stable. Remember that with Keynes, investment is an autonomous variable, governed by long-period expectations and, as such, psychologically determined and highly volatile (Keynes, 1936, pp. 102–3).

The supermultiplier

The salient features of the supermultiplier relation may now be pictured. Large autonomous variables ($G^* + X^*$) obviously increase trend output Q^*, as emerges from equation (3.4) and Figure 3.3 (the output and employment trend is shifted upwards). The same is true for a large gross investment–output ratio (a^*): investment volumes are particularly large when great inventions are realized and become innovations.

Government expenditures may increase dramatically in great wars and with armament races. For example, it is well known that it was not President Roosevelt's New Deal that overcame the Great Depression of the 1930s, but

World War II. And it is evident that very successful exporters of manufactured products such as Germany, Japan and Switzerland enjoy permanently very high employment levels. On the other hand, trend output Q^* is negatively linked with a strong import dependence, as is reflected in a large import coefficient (b^*). Most importantly, an unequal income distribution, in principle, is associated with a lower level of output and employment: the spending power of the population is reduced and shows up in a low consumption–income ratio (c^*). The inverse long-period link between unequal distribution and output and employment is the crucial feature of the supermultiplier relation. This centrally important relationship between unequal distribution and involuntary unemployment represents, according to Schumpeter, the essence of the Keynesian revolution: '[The Keynesian doctrine] can easily be made to say both that "who tries to save destroys real capital" and that, via saving, "the unequal distribution of income is the ultimate cause of unemployment." This is what the Keynesian Revolution amounts to' (Schumpeter, 1946, p. 517).

Given these features of the supermultiplier, classical–Keynesian socioeconomic policies (incomes and employment policies in particular) are fundamentally about social improvement. Hence classical–Keynesian political economy is, in a Keynesian vein, a moral science essentially. This stands in sharp contrast to the natural sciences flavour of neoclassical mainstream economics.

Cycles and trend

Business cycles (that is, cyclical movements around the trend as shown in Figure 3.3) represent essentially an interaction between the investment behaviour of entrepreneurs and the socioeconomic system governing the institutional trend (Bortis, 1997, pp. 135–42, 204–20). Schumpeter (1939, Vol. 1, p. 213) puts the long-period Kondratiev cycles to the fore, around which shorter cycles, for example the Juglar cycles, are situated; the full Kondratiev cycle covers 50–60 years approximately (phases I–IV in Figure 3.3). How the Kondratiev cycle functions in principle can be explained most conveniently with the help of Figure 3.3. Here, trend output Q^* is governed by trend effective demand, represented by the right-hand side of equation (3.4), whilst realized output Q is determined by actually prevailing effective demand exhibited by the right-hand side of equation (3.5). This equation represents the pure theory of the business cycle. As such it is a variant of the supermultiplier relation in which both the realized investment–output ratio (a) and the realized consumption–income ratio (c) deviate from their trend values in the course of the cycle for reasons to be explained below:

$$Q = \frac{G^* + X^*}{(1 - c) + b^* - a} \tag{3.5}$$

In accordance with the conception of pure theory, the trend output (Q^*) (relation 3.4) is independent of realized output (Q) (relation 3.5), since all variables and parameters are abstract and set in logical time (Joan Robinson), not in historical time.

This would not be the case in applied theory set in historical time, where concretely existing magnitudes are being considered: here the realized variables would not be independent of the trend variables. For instance, if in a cyclical upswing improved production technologies are introduced, income distribution may get more unequal permanently: the wage sum would be reduced and profits would rise, since a higher output would be produced by the same labour force. The purchasing power of the population would decline and so would the demand for consumption goods. The reduced effective demand would push the trend downwards, increasing thereby long-period trend unemployment in the really existing economy under consideration.

The pure theory of the business cycle is based upon the interaction between the income effect of investment and the capacity effect of investment. The income effect consists in the post-Keynesian interaction between investment and profits, while the capacity effect is grounded upon the relation between productive capacity and the output related to it, and effective demand. The income effect starts from the fact that in the cyclical upswing the realized rate of growth (g) exceeds the trend growth rate (g^*) exhibited in the supermultiplier equation (3.4), implying that realized investment (I) and the realized investment–output ratio (a in equation 3.5) exceed trend investment (I^*) and the trend investment–output ratio (a^* in equation 3.4). This implies that in the price equation (3.3) the realized mark-up, associated with the realized rate of profits ($k(r)$), exceed the corresponding normal magnitudes ($k(r^*)$). However, higher realized profit rates (r) will bring about higher realized growth rates and correspondingly larger investment volumes.

A cumulative process based upon the interaction between the investment volume (I) and realized profits (P) is set into motion (I–P in Figure 3.3). Yet, high investment volumes lead on to an expansion of productive capacities. Realized output (Q) gradually rises above the trend output (Q^*) as governed by the institutional-technical system through the supermultiplier equation (3.4). An increasing GDP (Q) exerts a pressure on prices and profits. At $t = A$ in Figure 3.3, the slope of the Q curve equals that of the

Q^* curve, implying that a and c in relation (3.5) equal a^* and c^* in relation (3.4). As a result, realized profit equals the target rate ($r = r^*$). However, realized output Q is AC, thus exceeding effective demand AB. Given this, realized profits in equation (3.5) will now fall short of the normal magnitudes set forth by the supermultiplier equation (3.4). The cyclical movement now changes direction and the downswing is initiated; the interaction between investment and profits (I–P) now works in the reverse way in the direction of a slump. The economy will pick up again once realized output has declined to a level below the trend output long enough such that the realized profit rate exceeds the normal profit rate implied in the long-term supermultiplier equation. The financial sector may reinforce the cyclical movements. In the upswing new credits grow very rapidly; in the downswing a credit crunch may occur.

The role of the financial sector in a monetary production economy and international trade in a classical–Keynesian perspective would be other important issues to be investigated (on finance, see Bortis, 2013, pp. 346–52 and Bortis, 2015; on international trade, see Bortis, 2013, pp. 356–8). In fact, classical–Keynesian political economy enables us to address all-important problems of the modern world in a sensible way.

Neoclassical–Walrasian economics and classical–Keynesian political economy assessed

Which approach – neoclassical–Walrasian economics or classical–Keynesian political economy – is more plausible and hence fit for policy purpose? Theoretical critique and historical-empirical facts strongly speak in favour of Keynesian and classical–Keynesian political economy. On the theoretical side, the above-mentioned capital-theoretic debate (Harcourt, 1972) has shown that there are no well-behaved associations between factor prices and factor quantities in general, and specifically between rates of interest (profits) and quantities of capital. This implies that, in principle, and in the long run, investment cannot adjust to full employment saving; hence the market cannot produce any tendency towards full employment at all, even in principle. This is crucial, because principles have to be solid like a rock, otherwise theories are built upon sandy foundations. On the historical-empirical side the great crises of the 1930s and of 2007–08 may be explained convincingly by the classical–Keynesian supermultiplier and the cyclical fluctuations taking place around the long-period trend (Bortis, 1997, pp. 142–220). Given this, Joseph Schumpeter was quite wrong in arguing that Walras's general equilibrium model was the *Magna Carta* of economic theory.

Moreover, devastating internal critique of modern neoclassical economics is being put forward at present. Indeed, the eminent MIT economist Ricardo Caballero writes: 'I am almost certain that if the goal of macroeconomics is to provide formal frameworks to address real economics rather than purely literature-driven ones, we better start trying something new rather soon' (Caballero, 2010, p. 87). 'The root cause of the poor state of affairs in the field of macroeconomics lies in a fundamental tension in academic macroeconomics between the enormous complexity of its subject and the micro-theory-like precision to which we aspire' (p. 100). Caballero thus concludes that:

> The challenges are big, but macroeconomics can no longer continue playing internal games. The alternative of leaving all the important stuff to the 'policy'-types and informal commentators cannot be the right approach. I do not have the answer. But I suspect that whatever the solution ultimately is, we will accelerate our convergence to it, and reduce the damage we are doing along the transition, if we focus on reducing the extent of our pretense-of-knowledge syndrome. (Caballero, 2010, pp. 100–101)

To this rather pessimistic statement this chapter adds a touch of optimism. Indeed, classical–Keynesian long-period theory – that is, the theory of the long-period output and employment trend – and its implication for the theories of value and distribution, as well as the role of the financial sector in a monetary production economy (see Bortis, 1997, 2003, 2013, 2015), represent the starting point for building up an open-ended classical–Keynesian system of political economy. On the basis of this theoretical system most differing aspects of an evolving real world may be addressed in an altogether different way from the currently dominating mainstream economics. (See Box 3.1 for further reading on the history of economic theories.)

Box 3.1

FURTHER READING

The present chapter is intended to initiate thinking about the history of economic theories. It is all-important, however, to continuously deepen the knowledge in this field through further reading. This means getting acquainted with selected parts of primary literature (or secondary literature directly linked with primary literature), some of which are mentioned in the References section, and reading systematic presentations of the history of economic theories. Proceeding in this way leads on to an interaction between this chapter and further reading: on one hand, this chapter should enhance understanding of selected primary literature and books on the history of

➡

economic theories, and on the other hand, further reading will bring about deeper insights into this chapter. This difficult and time-consuming process should, ideally, result in what Keynes called 'the emancipation of the mind', that is, to think for oneself, independently, in order to be able to distil the most plausible pure economic theory, on the basis of which economic phenomena can be explained – employment and distribution, for example – through applied theory. And one should always bear in mind that solid pure economic theory is the basis for appropriate economic policies; for example, employment and distribution policies. Nothing is more practical than good (pure) theory (applied to specific policy problems), a well-known saying goes. David Ricardo masterfully applied this saying as a member of the British Parliament, and so did Maynard Keynes, in his capacity as a writer on policy issues and as a government adviser.

The first book to be proposed is the excellent introductory volume by Heinz Kurz:

Heinz Kurz, *Economic Thought: A Brief History*, New York: Columbia University Press, 2016, paperback 2017.

Three intermediate texts to be highly recommended are:

Phyllis Deane, *The Evolution of Economic Ideas*, Cambridge: Cambridge University Press, 1978.

Phyllis Deane, *The State and the Economic System: An Introduction to the History of Political Economy*, Oxford: Oxford University Press, 1989.

Ernesto Screpanti and Stefano Zamagni, *An Outline of the History of Economic Thought*, Oxford: Clarendon Press, 1993.

A masterful advanced text is:

Maurice Dobb, *Theories of Value and Distribution since Adam Smith: Ideology and Economic Theory*, Cambridge: Cambridge University Press, 1973.

Moreover, there is the great comprehensive volume by Joseph Schumpeter:

Joseph Schumpeter, *History of Economic Analysis*, London: Allen & Unwin, 1954.

NOTE

1 Also known as the Marshallian scissors.

REFERENCES

Bortis, H. (1997), *Institutions, Behaviour and Economic Theory: A Contribution to Classical–Keynesian Political Economy*, Cambridge, UK: Cambridge University Press.

Bortis, H. (2003), 'Keynes and the classics: notes on the monetary theory of production', in L.-P. Rochon and S. Rossi (eds), *Modern Theories of Money: The Nature and Role of Money in Capitalist Economies*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 411–75.

Bortis, H. (2012), 'Towards a synthesis in post-Keynesian economics in Luigi Pasinetti's contribution', in R. Arena and P.L. Porta (eds), *Structural Dynamics and Economic Growth*, Cambridge, UK: Cambridge University Press, pp. 145–81.

Bortis, H. (2013), 'Post-Keynesian principles and economic policies', in G.C. Harcourt and

P. Kriesler (eds), *The Oxford Handbook of Post-Keynesian Economics, Volume 2: Critiques and Methodology*, Oxford, UK and New York, USA: Oxford University Press, pp. 326–65.

Bortis, H. (2015), 'Capital mobility and natural resources dynamics: a classical-Keynesian perspective', in M. Baranzini, C. Rotondi and R. Scazzieri (eds), *Resources, Production and Structural Dynamics*, Cambridge, UK: Cambridge University Press, pp. 155–73.

Bortis, H. (2019), 'John Neville Keynes', in W. Dimand and H. Hagemann (eds), *The Elgar Companion to John Maynard Keynes*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 10–16.

Caballero, R. (2010), 'Macroeconomics after the crisis: time to deal with the pretense-of-knowledge syndrome', *Journal of Economic Perspectives*, **24** (4), 85–102.

Dobb, M. (1973), *Theories of Value and Distribution Since Adam Smith: Ideology and Economic Theory*, Cambridge, UK: Cambridge University Press.

Garegnani, P.A. (1983), 'Two routes to effective demand', in J.A. Kregel (ed.), *Distribution, Effective Demand and International Economic Relations*, London: Macmillan, pp. 69–80.

Gossen, H.H. (1854/1983), *Die Entwicklung der Gesetze des menschlichen Verkehrs und der daraus fließenden Regeln für menschliches Handeln* [*The Laws of Human Relations and the Rules of Human Action Derived Therefrom*], Boston, MA: MIT Press.

Harcourt, G.C. (1972), *Some Cambridge Controversies in the Theory of Capital*, Cambridge, UK: Cambridge University Press.

Keynes, J.M. (1926), *The End of Laissez-Faire*, London: Hogarth Press.

Keynes, J.M. (1930a/1971), *Treatise On Money, Volume I: The Pure Theory Of Money*, reprinted as Volume V in *The Collected Writings of John Maynard Keynes*, London, UK and New York, USA: Macmillan / Cambridge University Press.

Keynes, J.M. (1930b/1971), *Treatise On Money, Volume II: The Applied Theory Of Money*, reprinted as Volume VI in *The Collected Writings of John Maynard Keynes*, London, UK and New York, USA: Macmillan / Cambridge University Press.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Marshall, A. (1890/1920), *Principles of Economics*, London: Macmillan.

Marx, K. (1867–94/1973–74), *Das Kapital*, 3 vols, Berlin: Dietz-Verlag; first editions 1867, 1885 and 1894.

Meek, R.L. (1962), *The Economics of Physiocracy*, London: George Allen & Unwin.

Pasinetti, L.L. (1977), *Lectures on the Theory of Production*, London: Macmillan.

Quesnay, F. (1758), *Tableau économique*, Versailles: privately printed.

Ricardo, D. (1810), *The High Price of Bullion, a Proof of the Depreciation of Bank Notes*, London: J. Murray.

Ricardo, D. (1815), *An Essay on the Influence of a Low Price of Corn on the Profits of Stock*, London: J. Murray.

Ricardo, D. (1817), *Principles of Political Economy and Taxation*, 1st edition, London: John Murray.

Ricardo, D. (1821/1951), *On the Principles of Political Economy and Taxation*, 3rd edition, Cambridge, UK: Cambridge University Press.

Robinson, J. (1962), *Essays in the Theory of Economic Growth*, London: Macmillan.

Say, J.B. (1803), *Traité d'économie politique*, Paris: Déterville.

Schumpeter, J.A. (1939), *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process*, 2 vols, New York, USA and London, UK: McGraw-Hill.

Schumpeter, J.A. (1946), 'John Maynard Keynes, 1883–1946', *American Economic Review*, **36** (4), 495–518.

Schumpeter, J.A. (1954), *History of Economic Analysis*, London: Allen & Unwin.

Smith, A. (1759/1976a), *The Theory of Moral Sentiments*, Oxford: Clarendon Press.

Smith, A. (1776/1976b), *An Inquiry into the Nature and Causes of the Wealth of Nations*, 2 vols, Glasgow Edition.

Sraffa, P. (1960), *Production of Commodities by Means of Commodities*, Cambridge, UK: Cambridge University Press.

A PORTRAIT OF DAVID RICARDO (1772–1823)

David Ricardo was born into a wealthy Jewish merchant family. He attended elementary school and had some commercial training. At the age of 14 he started to work at the stock exchange. In 1792 his conservative father disinherited him because he had married a Christian and converted to Christianity. Being without means, he obtained some loans from friends, which he made skilful use of at the stock exchange. At the age of 25 he was said to be richer than his father. Given his comfortable material situation, Ricardo decided to partially retire from business activities to study natural sciences, mainly mathematics, physics and mineralogy.

In 1799 he got to know about Adam Smith's *Wealth of Nations*. Ricardo was deeply impressed and decided to devote his life to the study of political economy. For ten years (1799–1809) he only read and took notes. In 1810 he published his first and much-noticed work: *The High Price of Bullion, a Proof of the Depreciation of Bank Notes*. Later (around 1830) his quantity-theoretic views on money and prices were strongly criticized by the Banking School, which advanced a theory of endogenous money creation by banks.

In 1809 Ricardo became a Member of Parliament in the United Kingdom. He was very influential because his clear-cut theoretical vision of the functioning of the economy enabled him to argue simply and convincingly. In 1815 he published *An Essay on the Influence of a Low Price of Corn on the Profits on Stock*, followed in 1817 by the first edition of his fundamental *Principles of Political Economy and Taxation*.

Ricardo died at the age of 51 of an inflammation of the middle ear. His work represents the first logically consistent and complete system of political economy. He may be considered the founder of pure theory, which builds upon fundamental principles. For example, his pure theory of price rests on the labour value principle, his theory of income distribution upon the surplus principle.

Ricardo's principles *cum* pure theory method became all-important in the history of economic theory. Karl Marx took it up and put it in a very wide philosophical and historical context. Alfred Marshall, with the help of Neville Keynes, separated neoclassical pure theory of price, distribution and employment, erected upon the principle of supply and demand and the associated marginal principle, from historical realizations of prices, distributional outcomes and employment levels; here, competition would be imperfect due to the existence of monopolies, for example, or policy mistakes could be made, which could eventually put the theory into question. But this did not affect the principles embodied in pure theory. Given this, neoclassical pure theory, derived from Marshall's *Principles of Economics*, became a fortress that has resisted heavy attacks by Marxists, Keynesians and post-Keynesians (Bortis, 2019). This theoretical technique – pure *cum* applied theory – was put to use by Maynard Keynes in his *Treatise on Money* (1930, vol. I, *The Pure Theory of Money* and vol. II, *The Applied Theory of Money*) and in his *General Theory* (1936), where he distinguishes between:

> the logical theory of the multiplier, which holds good continuously, without time-lag, at all moments of time [the multiplier as a piece of pure theory is prior to its historical

➡

realizations in space and time], and [the applications of the multiplier principle, that is capturing] the consequences of an expansion in the capital-goods industries [real investment] which take gradual effect [in space], subject to time-lag and only after an interval [hence in historical time] (Keynes, 1936, pp. 122–5)

Hence David Ricardo is a giant in economic theory not only because of the content of his work, but also because of his pure-theory method.

EXAM QUESTIONS

True or false questions

1. Economics is about the rational behaviour of economic agents on the marketplace, which is to be coordinated by markets such that full employment obtains.
2. Political economy aims at understanding the functioning of the socioeconomic system in order to establish theories of employment, distribution, value, and others.
3. François Quesnay's *tableau économique* represents the social process of production and constitutes as such the theoretical founding piece of political economy.
4. In a monetary production economy, goods are always exchanged against goods.
5. The principle of effective demand asserts that unemployment is voluntary.
6. According to the surplus principle, profits are produced by labour.
7. Say's Law states that general overproduction and involuntary unemployment are possible.
8. According to classical–Keynesian political economy, profits are due to the productivity of capital.
9. According to Ricardo, the value of a product is governed by the direct and indirect socially necessary labour-time necessary to produce it.
10. Walras's general equilibrium model is the basic model of neoclassical economics.

Multiple choice questions

1. The supermultiplier is an elaborated form of the Keynesian investment multiplier stating how the output (and employment) trend is determined.
   a) According to the supermultiplier, the economy is set into motion by autonomous variables (trend government expenditures and exports), which bring about a cumulative process of production of both consumption and investment goods. The total output so determined governs the output and employment trend.
   b) According to the supermultiplier, economic activity is determined by the autonomous investment volume multiplied by the investment multiplier.
   c) According to the supermultiplier, economic activity is governed by the forces of supply and demand.
   d) The supermultiplier states that supply creates its own demand.
2. According to the supermultiplier the long-period investment volume is:
   a) governed by long-run expectations;
   b) demand determined by autonomous expenditures;
   c) determined on the market for new capital goods;
   d) somehow determined by Keynes's animal spirits.

3. In a perfectly competitive neoclassical economy, the condition 'price equal to marginal costs' for each firm implies that:
   a) output is determined by demand;
   b) output is not determined at a given market price;
   c) each firm produces a maximum output; therefore, output is supply determined;
   d) if the individual firm expands output, the price falls and each firm produces a minimum output.
4. Joseph Schumpeter admired Walras and called his general equilibrium model the *Magna Carta* of economic theory, because:
   a) interdependent markets only determine the quantities of an economic system;
   b) Walras established a sufficient number of equations to determine all the prices and quantities of an economic system;
   c) prices and quantities are determined on isolated markets;
   d) prices and quantities are in all cases determined by a bargaining process between economic agents.
5. Alfred Marshall rendered operable the complex Walrasian general equilibrium model of interdependent equations through:
   a) considering mark-up pricing within firms and industries;
   b) using the partial equilibrium method, that is, considering one market only; in partial equilibrium diagrams the principle of demand and supply emerges explicitly through the demand and supply curves;
   c) looking at bargaining contracts between individuals;
   d) considering sales of shops and department stores.
6. François Quesnay's fundamental *tableau économique*:
   a) sets out a linear view of production with industry and agriculture producing the social product;
   b) represents for the first time the social and circular process of production with industry and agriculture interacting to produce the social product;
   c) sets forth an economy in which industry alone produces the social product;
   d) implies an exchange economy in which goods are exchanged against goods.
7. The macroeconomic price equation implies that:
   a) the value of a product is determined by the conditions of production (various production coefficients);
   b) the value of a product, in principle, is governed by the direct and indirect labour time socially necessary to produce it;
   c) the money wage rate is not relevant in the determination of prices and hence the general price level;
   d) the value of a product depends upon the amount of utility it provides to consumers.
8. With classical–Keynesian political economy distribution is:
   a) regulated by the marginal productivity of the factors of production (labour, land and capital);
   b) fundamentally regulated by power relations and ought to be regulated by the principle of distributive justice;
   c) regulated by the law of demand and supply;
   d) basically a market issue.
9. In the political economy approach:
   a) competitive markets stand at the centre of considerations;
   b) the level of employment is governed by market forces, that is, supply and demand for labour;

c) the social process of production stands at the centre of considerations;
d) income distribution is governed by the supply and demand of the factors of production.

10. In neoclassical economics:
a) the social process of production is the main object of analysis;
b) the prices of goods and services and of the factors of production are governed by markets;
c) income distribution is regulated by power relations between capitalists and workers;
d) the level of employment is determined by effective demand.

4

Monetary economies of production

Louis-Philippe Rochon

OVERVIEW

This chapter:

- presents the neoclassical or mainstream view about the origins of money as linked to barter;
- explores the heterodox view of the creation and destruction of money through the existence of a monetary circuit of production;
- discusses the theory of the monetary circuit and the role of the banking system;
- describes the differences between a real and a monetary economy, and the relationship between bank credit, debt, money and production.

KEYWORDS

- **Barter theory of money:** According to neoclassical economists, the existence and creation of money are linked to barter and exchange.
- **Bursts of optimism and pessimism:** This expression describes the behaviour of firms or banks with respect to their expectations of aggregate demand in the unknown and unknowable future.
- **Hoarded savings:** The portion of savings that is not channelled through the financial market but represents the final debt of firms toward the banking system.
- **Monetary economy of production:** A money-using economy in which the emphasis is on the production of goods, financed by banks.
- **Reflux:** When money returns to firms through either the consumption of goods by households, or the purchase of private sector shares on financial markets.

Why are these topics important?

In the opening paragraph of *The General Theory of Employment, Interest and Money*, John Maynard Keynes (1936) warns us that applying policies arising from the teachings of neoclassical theory would be 'disastrous'. We have all seen the consequences of following monetary and fiscal policies derived from such an approach. Indeed, the economic and financial crisis that erupted in 2007–08 is the direct result of such policies and, as Keynes warned us, the consequences, which we are still feeling at the time of writing, have indeed been 'disastrous'.

Despite this global crisis, many economists still believe that neoclassical theory is fundamentally sound. This is indeed a strange thing, since a careful study of this theory reveals that it is incapable of predicting the possibility of a crisis: it is a theory of convergence and stability. As such, the theory breaks down and becomes irrelevant to understand the real world. While these same economists have claimed that there is no theoretical alternative (TINA), the purpose of this chapter is to show that in fact there is a coherent theory capable of explaining precisely how economies break down: the post-Keynesian theory of the monetary circuit.

One of the many failures of neoclassical theory is that it is detached from the real world, in particular because it gives no role to money. According to this theory, money is a 'veil' and is not allowed to interfere with the workings of the real economy. This approach, however, is clearly wrong: money plays an important role in our economic systems. In this sense, we must study what Keynes has called a 'monetary economy of production'. Money and banks are relevant, and within this 'monetary analysis', we can see the possibilities of crises arising.

The neoclassical/mainstream view

In neoclassical economics, there is no need for money to explain output, employment, consumption, wages, investment, growth, or even prices. You may wonder how one could discuss all these issues without money, yet this is exactly what neoclassical theory does, and when money is added to the story it is as a mere afterthought, a way of making the story seem more realistic. Ultimately, money plays but a minor role in neoclassical thinking (Box 4.1).

The origins of money in neoclassical analysis can be found in barter, a situation where two individuals exchange something they each own. There are many stories of barter, even today. For instance, we can imagine a carpenter

BOX 4.1

THE PLACE OF MONEY IN TEXTBOOKS

Students are encouraged to pick up any macroeconomics textbook, at any level, to see that money appears only much later in the book. This suggests that the real side of the economy (employment, demand, investment, growth) is independent from the monetary side. This book adopts a different approach: it begins with a discussion of money and the links between money and real variables, implying clearly that the real side and the monetary side of the economy are linked together: there cannot be a discussion of employment, investment, production and growth without first understanding the role of money and the banking system.

doing work for a dentist in exchange for a free dental examination. No money changes hands, but one commodity or service is simply traded for the other. Hence, for neoclassical economists, this was how markets operated before the existence of money: individuals simply traded with each other.

For instance, imagine a tailor wants a new pair of shoes. He would offer shirts he made to the cordwainer (the person who makes shoes), agree on a price, say three shirts for one pair of shoes, and make the trade. The price of a shirt in terms of shoes is called a relative price. Of course, a shirt would carry many prices, as many in fact as there are goods to trade. For instance, a shirt could be worth two pairs of socks, or one belt. A coat would be more expensive and be worth ten shirts. Relative prices are therefore the price of a good expressed in terms of different goods. This means that a price can be expressed in non-monetary terms: one shirt equals one belt. No need for money to explain a price. If the cordwainer is prepared to accept three shirts, a trade then takes place and both parties obtain what they want. But imagine what happens if the cordwainer is not interested in acquiring shirts; instead, he is looking for a pair of gloves. The tailor must then approach a glover, and propose to trade with him at a price that they agree upon. Once the trade takes place and the tailor has the gloves, he can then go back to the cordwainer and trade the gloves for shoes. Now imagine the glover is not interested in shirts, but wants a hat. The tailor must now approach the hat maker to trade shirts for hats at an agreed-upon price. Once the tailor has a hat, he can then approach the glover, trade the hat for the gloves, with which the tailor can finally approach the cordwainer and get the pair of shoes, which is all he wanted to begin with.

As one can imagine, this process can be difficult, time-consuming and costly: the tailor would have to invest considerable time in order to go

BOX 4.2

THE DOUBLE COINCIDENCE OF WANTS

In economics, the double coincidence of wants is a phenomenon that occurs when two individuals have each something the other wants. In the example used here, this would mean that the tailor wants to exchange shirts for shoes, and the cordwainer wants to trade shoes for shirts. In the neoclassical world, if this happened all the time, then there would be no need for money: markets would be perfectly organized and trade would be very smooth. Yet, they argue, the double coincidence of wants rarely occurred, thereby creating the need for money.

through this long process. If only there was some good that could be acceptable by everyone for trade, then the tailor would avoid all these additional exchanges.

This is where the invention of money enters neoclassical analysis: to make exchange between individuals easier. There would be no need to exchange several goods to finally end up with the goods you were originally seeking. In this sense, money was invented to make trade easier and to solve the problem of what economists call the 'double coincidence of wants' (Box 4.2).

Stanley Jevons (Box 4.3) first used this expression in the opening chapter of his book, *Money and the Mechanism of Exchange* (1875, p. 4):

> the first difficulty in barter is to find two persons whose disposable possessions mutually suit each other's wants. There may be many people wanting, and many possessing those things wanted; but to allow of an act of barter there must be a double coincidence, which will rarely happen.

BOX 4.3

STANLEY JEVONS

Stanley Jevons (1835–82) was a British economist largely considered one of the founding fathers of neoclassical economics, and is associated with the beginning of the mathematization of economics. Jevons is credited with the utility theory of value. By moving the analysis of macroeconomics away from social classes under classical economists like Karl Marx, Adam Smith and David Ricardo, and toward the marginal behaviour of individuals, Jevons is credited with what is called the 'Marginalist Revolution' (see Chapter 3).

According to this view, money would somehow emerge spontaneously to resolve the problem posed by the double coincidence of wants, and would take the form of a good or commodity that was highly regarded as tradable, such as precious metals (for instance, gold). Money's primary purpose therefore is to serve as a medium of exchange: money facilitates the exchange of goods between individuals. But note that in neoclassical theory there is no need for money: production takes place (the tailor is able to make his shirts, the cordwainer is able to make his shoes), and trade does occur. Yet, without money, it is just a little bit more difficult, and there are costs involved. In this sense, money is a 'veil': all activities take place behind the veil of money, and money is not allowed to influence anything. For instance, access to money will not encourage the tailor to make more shirts.

This story of the relationship between money and barter seems plausible: after all, we may have all perhaps bartered one service for another with a neighbour, for instance, so the story must be true; it makes sense. In fact, those who have studied the phenomenon tell a different story.

There is no doubt that barter existed in earlier societies; it is not difficult to imagine a farmer trading chickens for sandals. But the question is whether barter was pervasive and dominated early societies, and then whether money was really invented to make trade easier. American economist Paul Samuelson, a Nobel laureate, expressed this view very clearly: 'Even in the most advanced industrial economies, if we strip exchange down to its barest essentials and peel off the obscuring layer of money, we find that trade between individuals or nations largely boils down to barter' (Samuelson, 1973, p. 55).

While the economists' story seems plausible, anthropologists and historians, however – scholars who have actually studied the phenomenon – are not at all convinced. Instead, they argue that barter, while it did exist among some individuals, was not widespread and certainly did not dominate primitive societies. In fact, it may even not have had an economic purpose.

For instance, anthropologist Caroline Humphrey (1985, p. 49), of the University of Cambridge in the United Kingdom, argues rather convincingly that 'we know from the accumulated evidence of ethnography that barter was indeed very rare as a system dominating primitive economies'. Referring to barter as an 'imagined state', she then adds that '[n]o example of a barter economy, pure and simple, has ever been described, let alone the emergence from it of money; all available ethnography suggests that there never has been such a thing' (Humphrey, 1985, p. 48).

BOX 4.4

WHY BARTER?

A more interesting question then becomes: why do economists believe in the barter story when historians and anthropologists who have studied it claim that it was never a widespread form of exchanges? This is an interesting question to which there is no easy answer. But one possible answer is that it fits the economist's story, it allows them to set aside the messy story of what the existence of money can do: generate a crisis. Instead, in the face of a strong lack of evidence in favour of barter, economists seem to offer nothing more than to ask us to have faith in its existence.

Yet, despite the wealth of historical and anthropological evidence, mainstream economists have steadfastly ignored this evidence, and economics textbooks today still tell students the story of money, trade and the double coincidence of wants (Box 4.4).

But, above all, what this mainstream approach shows is that the theory of exchange is wholly independent of the theory of money, its creation or existence. Indeed, the cordwainer was able to fabricate shoes, and the tailor was able to produce shirts and then trade them even in the absence of money. They could even have hired some helpers (employment) in their workshops and paid them in shirts and shoes (wages), with which these workers could then go out to barter for goods that they want.

As one can see, in this neoclassical/mainstream view, money is added to the story much later, seemingly like a mere afterthought; money is not allowed to interfere with trade, it simply makes things easier. It is in this sense that neoclassical economists argue that money is neutral: it has no impact on the real economy. This is what Joseph Schumpeter called 'real analysis':

> Real analysis proceeds from the principle that all the essential phenomena of economic life are capable of being described in terms of goods and services, of decisions about them, and of relations between them. Money enters the picture only in the modest role of a technical device that has been adopted in order to facilitate transactions . . . So long as it functions normally, it does not affect the economic process, which behaves in the same way as it would in a barter economy; this is essentially what the concept of neutral money implies. (Schumpeter, 1954, pp. 277–8)

As most undergraduate students learn in studying economics, it is somehow possible to learn about employment, wages, investment, output and

economic growth without once referring to the existence of money or the influence of finance. Keynes summarized well the mainstream view:

> The distinction which is normally made between a barter economy and a monetary economy depends upon the employment of money as a convenient means of effecting exchanges – as an instrument of great convenience, but transitory and neutral in its effect. It is regarded as a mere link between cloth and wheat, or between the day's labour spent on building the canoe and the day's labour spent on harvesting the crop. It is not supposed to affect the essential nature of the transaction from being, in the minds of those making it, one between real things, or to modify the motives and decisions of the parties to it. (Keynes, 1973a, p. 408)

Money, that is to say, is employed, but is treated as being in some sense neutral.

But how can a system where money and finance are so crucial be excluded from the analysis of economics? Is this approach a realistic explanation of the real world we live in?

The heterodox view

The previous section concentrated on the mainstream view of the economic system and showed two important elements: (1) the emphasis was on exchange, not production; and (2) money was invented in order to make barter or trade easier. But a capitalist economy is, by definition, a money-using economy or, more precisely, a monetary economy of production. As such, money must be at the heart of the theory of output and economic growth, otherwise what we are describing is not really a capitalist system.

The heterodox view takes as fundamental the role of money and finance. In fact, in direct contrast to the mainstream view, heterodox economists, and post-Keynesians in particular, consider that it is impossible to explain employment, investment, wages, prices, output, production and economic growth without first and foremost understanding the role that money plays in the story. In this sense, the theories of output and economic growth are linked to the theory of money; unlike barter, it is impossible to talk about economics without referring to money. This is one of the most fundamental ideas of the heterodox approach.

While there are various heterodox approaches in economics, this section will concentrate on the contributions of an approach labelled the theory of the monetary circuit, which is a good summary of the heterodox approach

in general. It is a 'general theory' of a monetary economy of production that owes much of its central tenets to the views of John Maynard Keynes. Let us expand on this.

The theory of the monetary circuit

The approach known as the theory of the monetary circuit, or circuit theory, was pioneered by a number of economists, in both France and Italy, in the 1960s. Among the early proponents of this approach, three in particular stand out: Alain Parguez in France, Augusto Graziani in Italy, and Bernard Schmitt in both France and Switzerland (see the portraits of Alain Parguez in Chapter 5 and Bernard Schmitt in Chapter 11). All proposed elements of a common idea, and while there exist important differences between their analyses, it is possible to identify a general approach, which has been termed the 'monetary circuit'.

Two fundamental ideas characterize this approach. First, as explained below, the theory of output and the theory of money are linked. This is the essence of what Joseph Schumpeter called a 'monetary economy', as opposed to a 'real economy' that best describes the exchange economy of neoclassical theory as discussed above. Rather, a monetary economy, according to Schumpeter (1954, p. 278), 'introduces the element of money on the very ground floor of our analytic structure and abandons the idea that all essential features of our economic life can be represented by a barter-economy model'. This is what was stated at the beginning of this section: it is impossible to discuss employment and economic growth without first understanding the role of money.

Second, there is a clear sequence of events (sequential analysis) that must be respected: events occur in time. For instance, a firm cannot produce anything before it hires workers and purchases machines, which in turn cannot happen until the firm is able to secure some sort of financing. Right away, you can imagine the possibility of a crisis: if a firm cannot secure financing, it cannot hire workers and machines. Production does not take place. In the words of Keynes (1973a, p. 408), '[i]n my opinion the main reason why the problem of crises is unsolved, or at any rate why this theory is so unsatisfactory, is to be found in the lack of what might be termed a *monetary theory of production*'.

Keynes wanted to write about the way 'the economy in which we live actually works' (Keynes, 1936, p. 12). In a chapter contained within the first proofs of *The General Theory*, but which did not make it to the print, Keynes (1979, pp. 67–8) describes his endeavour as follows: 'It is to the theory of a generalised monetary economy, i.e. of an economy in which, through the fault or the

inaction or the impotence of the monetary authority . . . that this book will attempt to make a contribution'. Elsewhere, Keynes (1973a, p. 411) makes a similar statement:

> Accordingly I believe that the next task is to work out in some detail a monetary theory of production, to supplement the real-exchange theories which we already possess. At any rate that is the task on which I am now occupying myself, in some confidence that I am not wasting my time.

This task is to integrate money within the analysis of production, at the very beginning of the story. In this light, the Canadian economist Marc Lavoie (1984, p. 773), a leading heterodox economist, reminds us the importance of integrating money from the very beginning of the analysis: 'The injection of money in the economic system must not be done when output is already specified, as in the exchange economy . . . but rather must be introduced as part of the production process'.

Thus, as the name suggests, the theory of the monetary circuit involves the importance of money and production, within a realistic view of our contemporary economies, and as part of a circuit it shows how money is first created, then circulated, and then finally destroyed.

Core assumptions

Let us begin our analysis of the monetary circuit with a short list of five core assumptions. These are as follows:

1. The economy is best described as happening in historical time, meaning that events occur in time and not all at once; events are irreversible. For instance, workers cannot consume before they have found employment that gives them an income from which they can spend and save; corporations cannot hire workers before they have secured the necessary funding to pay for production and wages; once a firm has purchased a capital good (a machine), it cannot reverse its decision. In other words, the present is characterized by decisions taken in the past that cannot be (easily) undone. Similarly, while the past is known and cannot be changed, the future is unknown and, in many respects, unknowable (Box 4.5). This makes decisions regarding investment, for instance, difficult.
2. Contrary to neoclassical theory, which puts the individual at the heart of the analysis, heterodox economists adopt a class-based approach, reminiscent of the classical school. These macro-groups (Graziani, 2003) are:

BOX 4.5

UNCERTAINTY

Uncertainty is at the heart of the heterodox approach. By 'uncertainty' we mean a situation when the outcome of an action is not only unknown, but we cannot predict its probability. Take, for instance, a game of cards. There is nothing uncertain about it. We know with certainty the probability of, say, getting a queen (there are four queens in a deck of cards). We also know the probability of getting red in roulette. In that sense, betting everything you have is risky, but not uncertain: you have a 47.37 per cent chance of winning in American roulette. But knowing the rate of interest set by the central bank one year from today is impossible to know; this is how the future is unknown and unknowable.

workers (who are paid a money-wage), firms or non-financial corporations, banks and the government. A fifth group can be added to this analysis, that is, the rest of the world. The emphasis therefore is on how these macro-groups interact with one another, and the possible conflict that can arise from this interaction.

3. Privately owned firms are also divided into two subsectors, namely, firms that produce consumption goods (destined to be sold to workers), and firms that produce investment goods (for instance, machinery; destined to be sold to other firms).
4. The banking system is at the heart not only of the production process but also of the creation of money through the supply of bank loans (credit); this is the essence of what is called the endogeneity of money: banks make production possible, and money is the result of bank credit.
5. The analysis is made in terms of a period of production, which is defined as the time between the creation and destruction of money. This is not in calendar years, but rather the logical flow of money.

To further analyse the theory of the monetary circuit, let us consider five specific stages of the circuit (see Rochon, 1999, for further analysis).

Stage one: the planning of production

The circuit begins when non-banking firms plan their production levels, which are based on their expectations of aggregate demand in the near future. Being expectations, firms operate in uncertainty. In this sense, the supply of goods being produced is dependent on what firms believe the level of demand will be in the not-so-distant future; supply adapts to expected demand.

From this, two conclusions can be drawn. First, this is the rejection of Say's Law, which is at the core of neoclassical theory. However, heterodox economists believe that it is supply that adjusts to demand: if firms expect demand to increase, they will increase their level of production, or decrease it if they expect demand to weaken.

Second, in a world of uncertainty, expectations can be wrong: firms decide on production levels based on their expectations of demand, which can be easily frustrated. Changing perceptions of the future levels of aggregate demand can lead firms to change their production plans.

However, once production levels have been determined, firms are in a position to make a number of additional important decisions; for instance, how much of their existing productive capacity will be utilized in the production process, how much labour to hire, and the price of their product.

Firms typically never produce at 100 per cent capacity (see Chapter 9); the degree of capacity utilization will be less than full. This gives firms the ability to increase production in case demand becomes greater than what the firm anticipated. Imagine, for instance, the production of a new automobile that suddenly becomes very popular. In order to respond to the increased demand, at least in the short run, a firm will increase production by increasing its rate of capacity utilization. If demand falls, for instance because of a recession, firms will be able to respond by lowering their degree of capacity utilization.

Firms will then have to decide on how much labour to hire, given their expectations of demand, and the degree to which they will utilize their productive capacity. Given the wage rate, the firm will then know not only the level of labour, but also its wage bill, which is defined as the wage rate times the level of labour.

Finally, firms will be able to set the price of their product or service. This price will be based on the costs of producing the goods, over which they will add some mark-up based on their desired rate of profitability. For instance, if it costs them US$100 to produce a single unit, they may add a mark-up of, say, 25 per cent, thereby setting the unit price at US$125.

Stage two: bank credit and the creation of money

The next phase of the circuit will move from the planning of production to the actual production of goods, with its own challenges: since firms have not

BOX 4.6

SAY'S LAW

In economics, Say's Law stipulates that the supply of goods creates its own demand, thereby assigning a purely passive role to aggregate demand. In his 1803 book, *A Treatise on Political Economy* (originally published in French), Jean-Baptiste Say wrote: 'A product is no sooner created than it, from that instant, affords a market for other products to the full extent of its own value' (Say, 1803/1834, p. 138). In this case, therefore, there can never be a shortage of demand. For post-Keynesians, however, aggregate demand is the driving force of economic activity, thereby invalidating Say's Law.

sold anything yet, they typically do not have the necessary funds to start production. Post-Keynesians specifically reject the idea that pre-existing savings finance production. This idea, known as Say's Law, is a remnant of real analysis and neoclassical economics (Box 4.6). But when you think in terms of historical time, where events do not happen all at once but must follow the logic of the passage of time, then we must ask: how can firms have pre-existing savings if there are yet goods to be sold? As a result, out of necessary logic, we must reject Say's Law and ask: where do firms obtain the funds to begin the production process?

The obvious answer in a modern economy is that firms must turn to banks in order to secure the proper funding to cover their costs of production (or at least part of these costs): production is financed by bank credit.

This is an important argument: firms do not have access to prior funds or savings at this early stage. This is why bank credit and the existence of banks are such a crucial component of the monetary circuit: in order to produce, firms must borrow from banks and accept getting into debt. As Seccareccia (1988, p. 51) writes, production is 'a process of debt formation'. Indeed, there is a very specific relationship between the financial needs of production, bank credit and the creation of private sector debt. As will be seen below, we can add to this list the link with the creation of money proper.

But to secure funds from the bank, the bank must deem the borrowers creditworthy; if they are not, the bank will refuse to lend them the funds. Typically, this means that the bank must be satisfied that the borrower, in this case the firm, will be able to reimburse its loan in the future with interest; an argument developed in the next section.

Once the bank is satisfied with the creditworthiness of the borrower, it will lend the necessary funds to begin production. This usually takes the form of a line of credit, which the firm draws upon to pay wages and cover other costs related to the production of its goods.

Once wages are paid, money is then deposited into the bank accounts of workers, at which point money is created: the process of money creation is linked to the payment of wages, which in turn is linked to the firm's plans to produce and secure the necessary debt to do so. Notice how fundamentally linked the theory of money creation is to the theory of output and production, as discussed previously. This is a far cry from the exchange economy of neoclassical economics: in the monetary circuit, money is created when firms agree to get into debt with regard to the banks, and when wages are paid to workers. Once wages are paid, workers release these funds into active circulation when they consume, a point which will be returned to below.

But so far, from the above discussion, two conclusions can be reached. First, the supply of bank loans is made at the initiative of the borrower. Banks cannot lend if there is no demand for loans. This is the embodiment of the old adage 'you can bring a horse to water but you can't make him drink'. It is in this sense that the supply of loans is demand-determined. This is another example of the rejection of Say's Law: it is the demand for loans that creates the supply, as long as those who are demanding credit are deemed creditworthy. Second, these ideas are further developed in the next chapter, but one important conclusion here is that banks are not constrained in their ability to lend, except by the demand from creditworthy or good borrowers. This conclusion was reached several decades ago by British economist Joan Robinson. As she wrote, 'the amount of advances the banks can make is limited by the demand from good borrowers' (Robinson, 1952, p. 29).

Stage three: the bank's decision to lend

Let us revisit in some detail the decision of the bank to grant credit and the criteria it follows to make this decision. As stated above, banks are only limited by the demand from 'good' borrowers. But how do banks determine who is a 'good' borrower and who is not? Because we live in an uncertain world, the future is by definition unknown. In deciding to give credit to a firm, a bank faces two sources of uncertainty regarding the firm's ability to reimburse its loan in the future, what Rochon (2006) has called 'micro-uncertainty' and 'macro-uncertainty'.

Micro-uncertainty is the bank's evaluation of the competence of the firm's management, and whether the firm itself has a good record at reimbursing its past debt obligations. Does it have good leadership, a competent chief executive officer (CEO), a good track record, and did it produce a good market study? Banks will typically consider a number of criteria in order to reach a decision, such as past relationship with the firm, the firm's net worth and collateral, debt/equity ratio as well as other financial ratios. In other words, regardless of the state of the economy, this is all about the firm itself.

This was well summarized by Barker and Lafleur (1994, p. 83) when they argued that the 'role of banks depends largely on information systems that allow banks to determine the solvency of their customers. Their success will depend heavily on the intuition of their credit officers and their ability to identify the capacity and willingness of borrowers to repay loans'.

But even if the firm is deemed competent at the micro-level, there is another source of uncertainty: the bank's expectations of the macro-environment in the near future. Does the bank forecast a period of strong economic growth or a recession? This is important: if the bank believes a recession is forthcoming, characterized with a general decrease in income, then this will make it more difficult for a firm to sell its products and reimburse the bank. On the other hand, if the bank is anticipating strong economic growth, and therefore a growth in future incomes, then this is interpreted to mean that firms will benefit from this growth, and be able to sell more of their products: increased sales, increased revenues, and thereby increased ability to reimburse bank debt.

The question, of course, is: how do banks measure their expectations of aggregate demand? This is something no one can really do with any measure of accuracy. Data on variables such as unemployment, or sales, or any other relevant variables, often contradict each other and leave banks (and firms) with great uncertainty.

In the end, there are two sets of criteria, each aimed at two different objectives: on the one hand, to identify the competence of the firm's management, and, on the other hand, to identify the capacity, so to speak, of the economy as a whole. What is clear is that banks will lend only to those firms that meet the bank's strict lending criteria. As a result, as Keynes (1930/1971, p. 327) once said, there will always be a 'fringe of unsatisfied borrowers'.

The two sets of criteria will play a very different role in the bank's decision. Macro-uncertainty will determine the minimum criteria all firms must meet

BOX 4.7

THE ZERO LOWER BOUND FOR NOMINAL INTEREST RATES

There is an important lesson to be learned regarding the period of near zero nominal interest rates in many countries since the eruption of the global financial crisis in 2007–08. Why is it that at historically low rates of interest, banks were not lending more? The analysis above offers a clear explanation: despite low interest rates, banks were still very pessimistic about firms' abilities to pay back potential loans, given the banks' pessimism regarding the future of aggregate demand. In an era of secular stagnation, with an expected long period of low economic growth, one can expect banks to remain sceptical.

in order to get a loan. Moreover, these criteria will change with the bank's expectations of the future levels of aggregate demand. For instance, if the bank becomes more optimistic about the future, it will 'lower the bar', so to speak, making it easier for firms to qualify for a loan. This is because in a growing economy banks are more optimistic of a firm's ability to generate income. But if banks become more pessimistic, for instance if a bank expects a recession, then it will raise its criteria, thereby making it more difficult for all firms to qualify for a loan (Box 4.7).

Irrespective of these minimum criteria, if a firm meets them, it will be granted a loan (Le Bourva, 1992). The immediate conclusion we reach is that a bank is never constrained in its ability to grant loans. It is only constrained by the number of creditworthy borrowers (see Lavoie, 2014).

Now, what about the criteria used to deal with the micro-uncertainty? These will be used to identify the firm's 'degree of creditworthiness', or how robust firms are. If the firm already meets the basic minimum level of creditworthiness, how much more creditworthy is it? This will then be used to determine the rate of interest that will be applied to the firm's borrowing needs. If it meets the minimum level but the bank still sees some uncertainty at the level of the firm, the bank may impose a higher rate of interest than another firm that has a more robust micro-evaluation. Hence, while the macro-uncertain environment will determine whether a firm gets a loan, the micro-uncertainty will determine the mark-up over the minimum lending rate. The higher the micro-uncertainty, the higher the mark-up and hence the higher the loan rate.

Stage four: the reflux principle and the destruction of money

So far, production has been financed by bank loans, firms have got into debt, and wages have been paid. Firms must now sell their wares, generate revenue and reimburse the bank. How then are firms able to recuperate their production costs?

This is done largely during the process of consumption, when households spend their income on purchasing produced goods. As households use their income to buy consumer goods, money then flows back to the firms. Naturally, the more households spend, the more money finds itself back in the bank accounts of the firms: the higher the aggregate demand, the higher the ability for firms to recoup part of their initial costs.

At this point, one can see a clear circuit of money: money first goes out with the payment of wages, then flows back with the act of consumption, which becomes a source of revenue for firms. As Parguez (1997, p. 5) wrote, '[t]his . . . stage depicts the paramount characteristic of the capitalist economy: firms must be able to recoup money from the sale of their output'. Le Bourva (1992, p. 454) has called this the 'alternating movements of creation and cancellation of money'.

The act of money returning to the firms is known in the monetary circuit approach as the 'reflux phase'. With this revenue, firms are able to pay back at least a part of their loans to banks, at which point money is destroyed (Figure 4.1).

The above description and Figure 4.1 are of course an oversimplification of the real world, which is more complex with loans to households, for instance, and a government sector that spends and taxes. But it nevertheless allows us to see how money is created, circulates and is destroyed, which is the essence of a monetary economy of production. But it also allows us to reach one important conclusion: the danger of excessive savings.

Figure 4.1 The creation and destruction of money

The above discussion omitted the existence of saving, and was based on the assumption that households would spend all their wages, which flow back to firms, thereby allowing them to reimburse banks. Yet, imagine what happens when households begin to save: firms are unable to capture the entirety of the costs of production and, as a result, are unable to reimburse the totality of their debt to the bank: too much savings can jeopardize the stability of the economy.

This is certainly an odd conclusion, but seen from the perspective of the logic of the monetary circuit, it is an inevitable result of the analysis. At the macro-level, therefore, the overall debt of the system toward the banking system will be equal to the savings of households; collectively, all firms pay out the total amount of their wage bill. The most they can recuperate is the wage bill, if households consume all of their income. But if they do not, whatever they save and do not consume will represent the outstanding debt of firms toward the banking system (Box 4.8).

The above discussion, though correct, is incomplete and requires us to discuss the nature of savings in more detail. This is because savings can take a few forms. More specifically, we can divide savings into two distinct categories, each with their respective implications. On the one hand, we have what we call 'hoarded savings', which are savings you keep, say, in your bank account. You can also call this 'liquid savings'; this is the traditional way many of us think about savings: what we keep for a rainy day.

On the other hand, we have 'financial savings', which are the savings we use to purchase financial assets on the stock market. These financial savings, when used to buy initial offerings of shares from private sector firms, will also flow back to firms. This suggests that firms are able to capture a portion of household savings, which they can use to reimburse a greater portion of their debt (Figure 4.2). This is a fundamental conclusion of the monetary circuit,

BOX 4.8

THE PARADOX OF THRIFT

For a single household, saving may be a good thing. It allows us to accumulate savings for a rainy day, for instance. In that sense, we are encouraged to save. But if everybody saves, then the economy may suffer greatly. This is called the paradox of thrift: at the micro level, saving may be a good thing, but at the macro level, it can destabilize the economy. In general terms, these paradoxes are called 'fallacies of composition'.

Figure 4.2 The creation and destruction of money with savings

one that did not escape Keynes's attention. For him, 'consumption is just as effective in liquidating the short-term finance as savings is' (Keynes, 1973b, p. 221).

Finally, recall that, earlier on, it was argued that money was created along with bank loans. As it is created, money takes the form of a flow: money flows through the economy. At the end of the circuit, however, when firms pay back their loans, some money will stay in bank accounts in the form of hoarded or liquid savings. Therefore, there will be an observable stock of money: money now takes the form of a stock. In this sense, money is both a flow and a stock, and this stock corresponds to the portion of their savings households desire to hold as a liquid asset. According to Sawyer (1996, p. 51), '[w]hether the money thereby created remains in existence depends on the demand for money as a stock'. Lavoie (1992, p. 156) concurs: 'There is no difference between the outstanding amount of loans and the stock of money.' This is another fundamental conclusion of the theory of the monetary circuit.

Stage five: the planning of investment

So far, this chapter has discussed the importance of bank credit in enabling the production process, with an existing stock of capital goods or machines. There is therefore a natural link between credit, debt, money and

production. But it has not yet discussed the role of investment, and how it is financed.

While the issue of how investment is financed is a matter of great debate, it is argued here that investment, like production, is financed largely from two sources: the retained earnings of firms as well as bank credit, but with a twist. Because investment concerns itself with the purchase of capital goods (machines, factories), which last many years, the loan is equally reimbursed over many years, or rather over many periods of production. Let us discuss this in some detail.

Let us first determine what may influence investment decisions of firms (see Chapter 9 for an analytical elaboration). To get an idea, let us consider the difference between production and investment. As stated above, production depends on the firm's expectations of sales proceeds in the near future, which is another way of saying expectations of aggregate demand. In turn, this will influence the level of capacity utilization of the existing stock of capital goods. In normal times, firms aim at producing at a normal level of capacity utilization, but they may produce beyond this level if they believe aggregate demand in the future to increase beyond the expected normal levels.

But what if firms expect these levels of aggregate demand to increase beyond the near future? In other words, what if they expect aggregate demand increases to be permanent? One solution is to permanently produce beyond normal levels of capacity utilization. Another solution is to increase its capital stock. This makes sense only if firms expect permanent increases in aggregate demand. This is because adding a machine or a factory implies increased productive capacity over the life of the capital good. A firm will not build a new plant only to leave it idle. So the addition to the capital stock means that the firm is prepared to produce more each year, presumably over the life of the capital good, of course with varying intensity.

As a result, it can be summarized that firms make investment decisions at the end of a period of production, as they revisit their production plans for the next period, based on their expectations of the future (see Chapter 9 for a thorough discussion of investment). And the decision to invest is a permanent one, in the sense that once an investment is made, the firm is left with the added capital good. It makes no sense for a firm to purchase a capital good and not use it, or to build a new factory and not use it. This is why investment is sensitive to changes in demand. This is a very different

conclusion from that reached by mainstream economists, who see investment dependent foremost on the rate of interest.

Some implications of the theory of the monetary circuit

The above discussion leads to some interesting conclusions, especially regarding the possibility of an economic crisis. There are two possible sources of crisis based on the above discussion. First, if banks become very pessimistic, it is clear that they will not be willing to lend. In other words, not many firms will be able to meet the criteria set by the bank. This is problematic. If this occurs, then many production plans will not be fulfilled, and unemployment will increase.

Take, for instance, the global financial crisis of 2007–08. Irrespective of the precise cause of it, it is clear that once the economy started to deteriorate, banks became increasingly pessimistic, and a sort of self-fulfilling prophecy set in: a collapse of income made banks pessimistic, which led them to cut loans, which depressed the economy even more.

Second, a crisis could occur if the amount of hoarded savings increases, as discussed above. Increases in hoarding or liquid savings have a negative impact on the economic system, as this frustrates firms' abilities to reimburse banks and extinguish their debt. But why would households want to increase their liquid savings? It may be the result of their increased uncertainty about the future: households may want to save for a rainy day. This translates into a decline in revenue for the firms, as households spend less on consumption goods and also on financial assets. As this occurs, firms will be less able to meet their contractual obligations with the banks. This may translate into an increase of the micro-uncertainty referred to earlier. In this regard, Seccareccia (1996, p. 16) explained that 'it is only when households choose to withhold their savings from the financial capital markets and seek to hold a significant proportion of their saving in the form of bank deposits that difficulties of reimbursement appear'.

The role of the state in the monetary circuit

The above discussion brings us to the importance of the state in the monetary circuit, and the role it can play in preventing a crisis (Figure 4.3).

So far, one issue that stands out is the importance of uncertainty (optimism or pessimism) for spending decisions. While the future is always by definition uncertain, the state has a role to play in mitigating its effects. Indeed,

Figure 4.3 The monetary circuit with the state

fiscal expenditures play a role at both the micro- and macro-level of uncertain environments.

In increasing its spending, the government transfers sums of money from its accounts (the public sector) to the accounts of both firms and households (the private sector). For instance, when building roads or bridges, the state transfers money to the firms that are involved in their construction. Also, we can imagine a number of programmes aimed at transferring sums to individuals, who then use these sums to increase their consumption. In doing so, this makes firms' revenues increase, and helps them in paying down their debt. In this sense, it makes them more creditworthy at the microeconomic level in the eyes of the banks: government spending can help in reducing

micro-uncertainty. Firms will find it easier to secure new loans for the following period of production.

Moreover, by increasing spending, the state also contributes to increasing aggregate demand at the macroeconomic level, thereby reducing macro-uncertainty. This will help banks to become more optimistic about the near future and hopefully extend more loans. Also, by reducing macro-uncertainty, firms will also become more optimistic and more willing to borrow and perhaps invest.

Conclusion

This chapter presented the theory of the monetary circuit, which describes the natural ebbs and flows of money, between when it is created and ultimately destroyed. This theory places the banking system at the core of its analysis, which finances both production and investment, each depending on expectations of the future. In this framework, uncertainty can play havoc with both activities since expectations can be easily frustrated.

Banks are also subject to this uncertainty and act following bursts of optimism and pessimism. This can be a source of crisis, if banks become too pessimistic and refuse to lend, or when potential borrowers are unwilling to borrow. In this sense, the theory of the monetary circuit is not only a theory of credit, debt, money and production, but also a theory of economic crises.

In the end, this theory stands in contrast to mainstream or neoclassical theory, where money is a mere afterthought, added simply to make barter and trade easier. In this sense, money can never be a source of crises. This chapter has shown what exactly Keynes (1936, p. 3) meant when he stated that policies derived from neoclassical economics can be 'disastrous if we attempt to apply it to the facts of experience'.

REFERENCES

Barker, W. and L.-R. Lafleur (1994), 'Business cycles and the credit-allocation process: an institutional perspective', in *Credit, Interest Rate Spreads and the Monetary Policy Transmissions Mechanism*, proceedings of a conference held at the Bank of Canada, Ottawa, November.

Graziani, A. (2003), *The Monetary Theory of Production*, Cambridge, UK: Cambridge University Press.

Hahn, F. (1983), *Money and Inflation*, Cambridge, MA: MIT Press.

Humphrey, C. (1985), 'Barter and economic disintegration', *Man*, **20** (1), 48–72.

Jevons, S. (1875), *Money and the Mechanism of Exchange*, New York: D. Appleton & Co.

Keynes, J.M. (1919), *The Economic Consequences of the Peace*, London: Macmillan.

Keynes, J.M. (1930/1971), *The Collected Writings of John Maynard Keynes, Volume VI: A Treatise on Money, Part II: The Applied Theory of Money*, London and Basingstoke, UK: Macmillan.

Keynes, J.M. (1931), *Essays in Persuasion*, London: Macmillan.

Keynes, J.M. (1933), *The Means to Prosperity*, London: Macmillan.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Keynes, J.M. (1973a), *The Collected Writings of John Maynard Keynes, Volume XIII: The General Theory and After: Part I: Preparation*, London and Basingstoke, UK: Macmillan.

Keynes, J.M. (1973b), *The Collected Writings of John Maynard Keynes, Volume XIV: The General Theory and After: Part II: Defence and Development*, London and Basingstoke, UK: Macmillan.

Keynes, J.M. (1979), *The Collected Writings of John Maynard Keynes, Volume XXIX: The General Theory: A Supplement*, London and Basingstoke, UK: Macmillan.

Lavoie, M. (1984), 'Un modèle post-Keynésien d'économie monétaire fondé sur la théorie du circuit' ['A post-Keynesian monetary economics model based on circuit theory'], *Economies et Sociétés*, **18** (2), 233–58.

Lavoie, M. (1992), *Foundations of Post-Keynesian Economic Analysis*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing.

Lavoie, M. (2014), *Post-Keynesian Economics: New Foundations*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Le Bourva, J. (1992), 'Money creation and credit multipliers', *Review of Political Economy*, **4** (4), 447–66.

Parguez, A. (1997), 'Government deficits within the monetary production economy or the tragedy of the race to balance budgets', University of Ottawa, mimeo.

Robinson, J. (1952), *The Rate of Interest and Other Essays*, London: Macmillan.

Rochon, L.-P. (1999), *Credit, Money and Production: An Alternative Post-Keynesian Approach*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Rochon, L.-P. (2006), 'Endogenous money, central banks and the banking system: Basil Moore and the supply of money', in M. Setterfield (ed.), *Complexity, Endogenous Money and Macroeconomic Theory: Essays in Honour of Basil J. Moore*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 220–43.

Samuelson P. (1973), *Economics*, 9th edition, New York: McGraw Hill.

Sawyer, M. (1996), 'Money, finance and interest rates: some post-Keynesian reflections', in P. Arestis (ed.), *Keynes, Money and the Open Economy: Essays in Honour of Paul Davidson, Volume I*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 50–67.

Say, J.B. (1803/1834), *A Treatise on Political Economy*, 6th edition, Philadelphia, PA: Grigg & Elliott.

Schmitt, B. (1975), *Monnaie, salaires et profits* [*Money, Wages and Profits*], Paris: Presses Universitaires de France.

Schumpeter, J.A. (1954), *History of Economic Analysis*, New York: Oxford University Press.

Seccareccia, M. (1988), 'Systemic viability and credit crunches: an examination of recent Canadian cyclical fluctuations', *Journal of Economic Issues*, **22** (1), 49–77.

Seccareccia, M. (1996), 'Post-Keynesian fundism and monetary circulation', in G. Deleplace and E.J. Nell (eds), *Money in Motion: The Post Keynesian and Circulation Approaches*, London, UK and New York, USA: Macmillan and St Martin's Press, pp. 400–16.

A PORTRAIT OF JOHN MAYNARD KEYNES (1883–1946)

Born in Cambridge, England on 5 June 1883, John Maynard Keynes, or simply Maynard to his closest friends, is considered the most influential economist of the twentieth century, and ranks among the greatest economists of all time. He is widely considered the founder of macroeconomics.

His most important book, *The General Theory of Employment, Interest and Money*, published in 1936, remains one of the most influential books in economics, and is still greatly debated today. The degree of influence of this book cannot be overestimated. In many ways, the entire development of macroeconomics since 1936 is in some way an attempt to either verify or contradict the contents of that book, which challenges the core neoclassical belief that markets are self-stabilizing and, if left on their own, would gravitate toward equilibrium, as well as the notion that unemployment was voluntary. Keynes argued that in a money-using economy in which uncertainty is a core feature, markets could break down, and could be subject to periods of prolonged recession, during which unemployment is actually involuntary, that is, independent of the people's willingness to work. Unemployment is not caused by wages being too high, but by the lack of aggregate demand. This view of markets convinced Keynes that governments have an important role to play in not only stabilizing markets, but in promoting economic growth through the use of countercyclical fiscal policy.

Throughout his life, Keynes held positions in both the academic and the private sectors. He started his career as a civil servant in 1906, working in the India Office, but returned to the United Kingdom in 1908, and began lecturing at the University of Cambridge in 1909, and in 1911 he was made editor of *The Economic Journal*.

At the beginning of World War I, he began working for the Treasury, and by the end of the war, in 1919, represented the United Kingdom at the Versailles Peace Conference. This is where Keynes's support for the government of the day ended, as he strongly objected to the terms of the Versailles Treaty, arguing that imposing reparations on Germany would be catastrophic. Keynes felt he had no other choice than to resign from the Treasury. He explained his objections to the Versailles Treaty in a famous and indeed influential book, *The Economic Consequences of the Peace* (1919). Keynes did not accept that Germany should be made to carry the burden of huge war reparations, because it would punish her citizens. In that sense, he rejected the notion that 'year by year Germany must be kept impoverished and her children starved and crippled', because the consequences would be 'vengeance, I dare predict' (p. 250).

With the start of the Great Depression in 1929, Keynes's ideas began to evolve considerably; in 1933, he published *The Means to Prosperity*, which contained some early ideas that would then reappear in 1936, in his *General Theory.* By now, Keynes was on board with the use of countercyclical fiscal policy, and this policy would begin to spread. In fact, one of Keynes's great accomplishments was to provide a theoretical justification for the use of fiscal policy in times of recession. As he writes: 'For Government borrowing of one kind or another is nature's remedy, so to speak, for preventing business losses from being,

➡

in so severe a slump as the present one, so great as to bring production altogether to a standstill' (Keynes, 1931, p. 161).

In 1944, Keynes returned to the service of the British government as one of its representative to the Bretton Woods conference, which resulted in the development of what is now known as the Bretton Woods system, and helped to create the World Bank and the International Monetary Fund.

In 1942, Keynes became a member of the House of Lords, as Baron Keynes of Tilton, in the County of Sussex. He died on 21 April 1946.

EXAM QUESTIONS

True or false questions

1. Barter was widespread and formed an organized market system.
2. Risk and uncertainty concern the same thing.
3. Historical time is described as events that happen according to a specific sequence in time.
4. The banking system is at the heart of the heterodox theory of production.
5. Heterodox economists adopt a class-based approach to explain how economies work.
6. In a world of uncertainty, expectations are fixed even in light of changes in expectations.
7. Post-Keynesianism is a modern revival of Say's Law.
8. When banks become optimistic, they lend less.
9. High household savings can jeopardize the stability of the economy.
10. Government spending can help in reducing micro-uncertainty.

Multiple choice questions

1. The 'double coincidence of wants' means:
   a) trade between individuals is facilitated by an acceptable medium of exchange;
   b) when two individuals want to trade, they each have what the other wants;
   c) in trade, coincidence refers to the needs of markets to expand simultaneously;
   d) coincidence occurs when the needs and wants of a single individual are met.
2. Uncertainty is when:
   a) consumers cannot decide on whether to buy a good or not;
   b) the past and the future are unknown;
   c) the future is unknown and unknowable;
   d) actions are risky and influence our decisions to act.
3. Say's Law stipulates that:
   a) the supply of goods is dependent on the profits of production;
   b) the supply of goods creates its own demand;
   c) the demand for goods creates its own supply;
   d) demand and supply of goods are independent from one another.
4. The process of money creation is linked to:
   a) the government's decision to increase the printing of money;
   b) the convertibility of paper money into coins;
   c) the consumption of finished goods;
   d) the payment of wages.

5. In the monetary circuit, the reflux refers to:
a) a return to old theories of money;
b) a flow of money back to firms;
c) the creation of debt;
d) the ability of banks to create money.

6. The state contributes:
a) to more instability given the existence of uncertainty;
b) to higher taxes;
c) to greater stability given the existence of uncertainty;
d) to higher levels of uncertainty.

7. Creditworthiness is defined as:
a) the ability of the firm to hire workers;
b) the ability of the firm to reimburse the initial bank loan;
c) the inability of the firm to produce goods over their initial target level;
d) the ability of the firm to reach a price level consistent with full employment.

8. The paradox of thrift suggests that:
a) the act of personal saving enhances society's wealth;
b) individual savings raise investment;
c) individual saving is consistent with Say's Law;
d) saving at the micro level (individuals) is detrimental to macroeconomic stability.

9. Investment depends on:
a) the saving levels of individuals;
b) the expectation of permanent increases in aggregate demand;
c) the ability of firms to increase the wage level;
d) Say's Law.

10. In the monetary circuit, two possible sources of crises are:
a) too much government spending and high taxes;
b) too much government spending and low savings;
c) pessimistic banks and high savings;
d) optimistic lending and too much government spending.

Part II

Money, banks and financial activities

5

Money and banking

Marc Lavoie and Mario Seccareccia

OVERVIEW

This chapter:

- presents the heterodox approach to money and banking and contrasts it with the mainstream;
- explains why mainstream economists view money as a commodity that takes on the role of medium of exchange, with banks being intermediaries between savers and investors; while heterodox economists view money as a means of payment resulting from a balance sheet operation within a creditor/debtor relation, with banks being creators of money to finance production;
- focuses on the importance of the creation and destruction of money by the banking system and on the crucial role played by the interbank market for funds and the payments and settlement system;
- points out that credit-money creation is demand-led and that the mainstream supply-determined perspective on bank lending is erroneous and leads to misguided policies such as the quantitative easing policies implemented in many countries during and after the global financial crisis of 2007–08 and during the COVID-19 crisis of 2020–21.

Readers will thereby understand why money is not a scarce commodity and why the banking sector can create credit-money whose only constraint is the demand for loans and the creditworthiness of borrowers. Readers will also understand why banks are the source of the finance that initiates the production process, while non-bank financial intermediaries play a role in bringing together savers and business enterprises that have already undertaken investment to address their final financing needs during the reflux phase of the circulatory process.

KEYWORDS

- **Commercial banking:** Financial institutions engaged in the business of financing, that is, in a process of making loans, and also accepting deposits whose overall effect is to create or destroy money.
- **Flux/reflux principle:** A basic principle in which bank credit advances constitute the flow while income receipts from these expenditures are the reflux of a corresponding amount that ought normally to permit the removal of the original loans from banks' balance sheets.
- **Means of payment:** In a non-barter system, a payment occurs using a third-party liability, namely that of the central bank or a private commercial bank, for final settlement by extinguishing counterparty debt of an equivalent amount.
- **Monetary circuit:** The circular process of advancing credit-money and then destroying an equivalent amount once the borrower is able to recapture the principal of a loan for reimbursement, leading to a closure of the circuit.
- **Payments and settlement systems:** National systems for the clearing/settlement of payments within the banking system, in which proper functioning of interbank lending/borrowing increases financial stability through enhanced financial market liquidity, together with a central bank as lender of last resort.

Why are these topics important?

Banks play an essential role in modern societies because our economies rely on credit for expansion and for redirecting production. Banks also play an important role because they are at the centre of the payments system, allowing economic agents to carry out transactions between each other in an efficient way. Banks are also an important cog in the financial system, as they facilitate speculative operations, a role which is not necessarily a positive one as it may generate instabilities. This is why banks need to be supervised to meet desirable social goals, either by the central bank or by some central regulatory agency. Several proposals are being put forward nowadays in an attempt to curb the excesses of the banking and financial system, so it is important to get the fundamentals right when it comes to the roles of banks and central banks.

The traditional mainstream view

All mainstream economics textbooks introduce money as Adam Smith once described it, in his celebrated 1776 opus, *The Wealth of Nations*, as the "universal instrument of commerce" (Smith, 1776/1937, p. 28), which was invented in order to facilitate exchange. While someone must have first

conceived and designed the wheels of a cart to make it easier to move goods and reduce transport costs as individuals sought to trade their commodity surpluses more efficiently, according to this view, so was money invented in order to grease the wheels of commerce and to engage with less effort in commodity exchange.

Hence, just as individuals at some moment in human history came to recognize the benefits of the division of labour and began to 'truck and barter', so it was, we are told, that they eventually began to use certain commodities, usually precious metals such as gold and silver, in their new role as money. Money supposedly emerged spontaneously from barter exchange because of these commodities' particular characteristics of divisibility, portability, fungibility, durability and relative scarcity, that allowed them to take on the role of medium of exchange, unit of account and medium of deferment of consumption.

As this tale of money emerging from barter exchange is normally told in mainstream textbooks, money's origin is explained simply as the natural outcome of private cost-minimizing behaviour that had nothing to do with the legal recognition and formal legal actions often taken by the state to ensure money's general acceptability.

This traditional perspective, therefore, rests on a particularly antiquated vision of money. Money is essentially conceived as a commodity, like grain, cowry shells or metals, whose metamorphosis into a medium of exchange catapulted this commodity money onto an otherwise pre-existing and privately organized natural barter system. From this, it follows that money's principal purpose was to make this market exchange merely more efficient, thereby surmounting the obstacle of the double coincidence of wants plaguing less efficient barter economies that preceded monetary exchange.

On the basis of this mainstream narrative on the origin of money and monetary exchange that is told repeatedly in economics textbooks, there is often also associated the tale of how banks, as particular institutions arising from this profit-seeking behaviour of individual economic agents, first made their appearance, and whose history is intertwined with that of money.

Historically, while quasi-banking-related activities of advancing simple credit appeared almost at the same time that humans began record-keeping, one of the tales of modern banking institutions as loan makers/deposit takers and issuers of banknotes goes as follows: banks emerged, we are told, very late after the Middle Ages owing primarily to the conduct of profit-seeking

goldsmiths, especially in seventeenth-century England. Banks appear in this traditional story in the following way. As these primitive societies progressed through market exchange, a portion of this overall stock of money in the form of precious metals changed hands at a certain annual turnover rate (or monetary velocity) to acquire and validate monetarily the flow of privately produced commodities. Within these commodity money economies, these transactions generated a flow of money income accruing to the various counterparties in the monetary exchange. Individual economic agents receiving these incomes faced the following options: a portion of this income flow of precious metals could be re-spent, thereby generating a series of consumption flows per period, while another portion could be saved or accumulated as liquid holdings, since other forms of financial assets had not yet appeared.

Banks appear as intermediaries in collecting a community's accumulated liquid savings for safekeeping, for instance, as represented in the textbooks by the stereotypical goldsmith bankers of seventeenth-century England. With time, instead of leaving these stocks of precious metals to sit idle and withholding them from circulation, these profit-seeking goldsmiths began to lend the portion of the community's stock of commodity money stored in their vaults. This saving would then be lent to those more enterprising individuals seeking to borrow money to undertake investment, by charging them interest. Since a portion of the investment expenditures would be returning to the same banks in the form of bank deposits, profit-maximizing banks would then re-lend this money once again, generating further loans in excess of the initial bank deposits. The effect would be to create bank money, this being the difference between the initial reserves of precious metals (or base money) and the total outstanding deposits, as banks progressively leveraged themselves in relation to their initial gold reserves, through a process traditionally referred to as 'fractional reserve banking'.

Ostensibly, the only constraint on this multiple expansion of bank money envisaged within this traditional framework was the desire by profit-seeking banks to hold idle reserves of this commodity money in their vaults exclusively for precautionary purposes, and the desire on the part of the public to hold some cash for day-to-day transactions purposes. Moreover, this leveraged banking system could only function as long as only a small portion of depositors withdrew their funds regularly and predictably for transactions needs. Otherwise, if depositors collectively sought to withdraw their gold all at once, as in times of financial panic, this banking 'house of cards' would collapse. There would just not be enough 'hard money' in the system, owing to the fact that the total money supply (that is, the coins in circulation and bank deposits) would actually be some multiple of the initial commodity money that had originally been

deposited for safekeeping. Moreover, whether individuals held bank deposits or whether they held other forms of bank liabilities, such as privately issued bank notes, the problem would be the same. Whether it is through deposit liabilities or private bank notes, as long as individuals were prepared to hold these two types of bank liabilities (and did not withdraw their precious metal deposits because of lack of confidence), this would allow a bank to grant loans in excess of the original commodity money deposits that initiated the process.

In this traditional story, while banks can create bank money as some multiple of the initial commodity money that was originally deposited in the goldsmiths' vaults and that is re-lent over and over through a circulatory process of deposit/loan expansion, banking institutions are conceived merely as depositories or storehouses of some pre-existing money that was deposited, say, for safekeeping (Realfonzo, 1998). Regardless of whether this particular tale of banking (resting on either commodity money or surrogates of such commodity money, such as central bank notes) may or may not reflect actual historical reality, this particular perspective on banks as storehouses, whose principal function is that of profit-maximizing intermediation between savers and investors, has changed little in modern times. Instead of resulting from the depositing of precious metals nowadays, we are told that it results from the depositing of exogenous base money initially created and issued by the central bank, which, through bank lending and subsequent deposit creation, leads to a multiple expansion of the money supply; a relation sometimes described as the base money multiplier. Mainstream theorists would argue, therefore, that 'deposits make loans'. Banks are conceived as passive deposit takers that serve the useful function of intermediaries, namely private institutions whose purpose was to transfer depositors' money (supplied by households that save) to creditworthy borrowers who would use those liquid funds for investment (traditionally business firms seeking credit advances). This is depicted in Figure 5.1, with investment being determined by the rate of interest in the market for loanable funds.

Figure 5.1 The mainstream conception of banks as intermediaries between savers and investors

Within this conception of the monetary system, households initiate a process whereby their initial savings are channelled to firms for productive investment through the intermediation role of the banking sector. Jakab and Kumhof (2015) present this intermediary view in an ironic way, saying that savers deposit previously accumulated gravel in banks, which in turn lend the gravel to entrepreneurs who wish to use it for productive purposes. As described in Figure 5.1, this ensuing business investment generates an income flow, a portion of which is held as savings. These savings then return to the banking sector, by starting up a new process as a portion of these savings are accumulated as bank deposits.

In their role as intermediaries between savers and investors, profit-maximizing banks make out loans to firms to finance their investment, which through the saving process are once again deposited and re-channelled through the banking sector to finance further loan expansion. This feedback process is then propelled forward until the possibility of further loan making from the initial injection of deposits has completely worked itself out. However, the system depicted in Figure 5.1, in its essentials, is a supply-determined system, whose growth is constrained by the reserves, say, of precious metals, or in modern times by the initial reserves of central bank money that was deposited in the banking system. Hence, for the system to expand, it requires the supply of base money to increase, which then allows the banking sector to serve its crucial allocative role of distributing loanable funds for productive investments. Moreover, whether this base money represents the stock of commodity money or of central bank-issued money is of no theoretical significance, since in either case it is the supply of these initial reserves, regardless of their precise forms, which moves the banking system forward in its role as intermediary between savers and investors.

The heterodox perspective

The heterodox theory of money and banking stands this mainstream perspective on the nature and origin of money and the functioning of the banking system somewhat on its head. The notion that money emerged from barter does not find strong basis in anthropological history. Credit/debit relations stipulated in a particular accounting unit (normally enforced by law or custom) pre-date the appearance of organized market exchange. When money did appear, mostly through the actions of the state, it took on the role of means of payment for settling debt obligations, especially tax liabilities (see Peacock, 2013). Within this perspective, and in contrast to the mainstream view, the particular characteristic of the commodity chosen as the unit of account was of little significance. What mattered was that money

was an abstract social unit, which was sanctioned by the legal apparatus of the state, and which would then become the means to discharge liability in a creditor–debtor relation. From this, it would ensue that money was not just a particular commodity with some special characteristic feature to facilitate exchange in the context of a natural, previously organized barter exchange system. Instead, the heterodox view suggests that money, as a means of cancelling debt, probably pre-dated organized market exchange itself. Emerging through the expenditures of the state, the legally sanctioned currency entering circulation eventually would not only assume the role of a means of payment in extinguishing debt obligations, but also the role of medium of exchange, and an evolving store of liquidity within organized markets.

Money, in the sense conceived by heterodox writers, appears in organized markets not as a commodity having some special intrinsic attribute as money. It appears, instead, as a third-party liability having only an extrinsic social value as legal tender, bestowed on it through the legal apparatus of the state either via its monopoly control over the central issuer of the currency (the mint) as during the Middle Ages, or through the state's monetary arm in modern times, namely its central bank (see Parguez and Seccareccia, 2000, p. 101). Banking institutions originated not primarily from some mistrustful group of profit-seeking goldsmiths who were lending out precious metals while pretending to be holding them in their vaults for safekeeping. Banking institutions existed even before the seventeenth century, and these institutions were slowly becoming specialized in the business of finance through double-entry bookkeeping, especially the financing of inventories for long-distance trade, because of the trust that they inspired through the public holding of their 'I owe you's' (IOUs), often through the tacit or direct support of the domestic authorities where they were based. In fact, already in twelfth-century Venice and later, many early commercial banks, such as the Sveriges Riksbank and the Bank of England in the seventeenth century, actually began as government debt agencies that issued debt certificates or promissory notes, which were then used by merchants and eventually the public as means of payment. However, by the eighteenth and nineteenth centuries, banks typically acquired from the government authorities a charter, which was a certificate or license authorizing the operation of a bank, whose business involved that of making loans and collecting deposits. Also, during that era, the bank charter normally gave the private institution the right to circulate its own privately issued bank notes denominated in the currency units established by the state. Because of their convenience, these private bank notes came to compete with the coins produced by the mint, until the mid-nineteenth century, when chequing facilities permitted chequable deposits to become more important than private bank notes in circulation.

However, because of public distrust in the viability of a payments system that can easily succumb to bank failures, at around the same time during the nineteenth century, governments started to assert monopoly control over the issue of paper currency notes so that, by the twentieth century, private bank notes virtually disappeared from circulation in Western countries.

With the decline of primary/agricultural activities and the rise of modern industrial production, banks extended their activities from the financing of inventories to that of financing short-term circulating capital requirements to facilitate the process of production itself. By the nineteenth century, bank credit played a central role in the financing of industrial production. Because of the fears of short-term withdrawals of bank deposits, banks were expected to behave prudently by financing the short-term circulating capital requirements of business enterprises in accordance with the 'real bills' doctrine.[1] Bank credit was not to be channelled towards the long-term funding of fixed capital investment, which instead ought to rely on business retained earnings or, through the issuing of securities with the deepening of financial markets, by capturing household saving.

To understand the traditional role of banks within this heterodox approach, let us begin with the fact that much like the state liability issued by central banks, the private banking sector as a whole could create credit-money at the stroke of a pen or *ex nihilo* (which is the Latin expression for 'out of nothing'). This is because, contrary to the mainstream tale of goldsmith banking, banks as a whole, and as long as they move closely in tandem in their lending, are not constrained by the amount of reserves arising from their deposit-taking activity. Within this perspective, it is actually loans that make deposits. Indeed, as soon as a bank makes out a loan to a creditworthy borrower, through double-entry bookkeeping, there will appear a counterparty deposit, which will initially appear in the private borrower's account (or in the case of an online credit this would happen instantly), which is then used to carry out the borrower's spending need. If the borrower were a firm, this credit money would go towards the compensation of workers and/or the purchasing of material inputs for the production process. In the archetypal version of bank financing production, this credit gives rise to a circulatory process, as shown in Figure 5.2.

Unlike the tale of goldsmith banking, this process of money creation is not driven by some initial deposits entering the banking sector. If such were the case, then where would the deposit first come from, unless it comes from some outside source, such as a government? Within the private banking sector, deposits can only appear when loans are made to either businesses

Figure 5.2 The heterodox conception of banks as creators of credit money within the framework of a rudimentary monetary circuit

or households. Instead, the initial injection of credit money created *ex nihilo* is sometimes described as the initial finance. The 'monetary flow' arrow in Figure 5.2 represents these initial credit advances to firms, which generate an 'income flow'. Households, on the receiving end of this income flow, would have a choice to allocate their income towards consumption or saving. In a world with no household saving, all the income is spent, which then allows firms to capture all of this income generated by the initial bank credit advances (as shown in the 'expenditure flow' arrow), thereby permitting firms to extinguish their debts vis-à-vis the banking sector.

In a world (say, of the nineteenth century) with non-existent (or very shallow) financial markets, the only possible saving is represented in Figure 5.2 in its most unsophisticated form of holdings of bank deposits. This is shown by the return arrow from the household sector to the banks (unless one also considers the option of hoarding bank notes under the mattress or in the cookie jar). As can be seen in Figure 5.2, with the leakage into bank deposits from the household sector, this withholding of consumption spending in liquid form as bank deposits, representing household liquidity preference, can short-circuit the flux/reflux process and prevent the business sector from extinguishing its overall debt to the banks. With a certain portion of the income flow not being spent, this would thrust banks into an uncomfortable intermediary role of 'deposits making loans'. Hence, it is only in this crisis state of incomplete closure of the monetary circuit, with household saving held in its most liquid form, that the causality between loans and deposits is reversed, since the desire to hold liquid deposits de facto forces banks to try to re-extend loans that cannot be fully reimbursed via the reflux. However, this is hardly the most realistic scenario, as shown in Figure 5.3, when there exist organized financial markets, reflecting a historically more sophisticated phase of financial deepening with a greater variety of portfolio choice.

Indeed, in the heterodox literature, there is an important distinction that is made between the 'initial' finance of the process of monetary and income expansion and the 'final' finance, which is associated with the reflux phase of this balance sheet circulatory process (see Graziani, 2003). In Figure 5.2, we have only considered the most elementary case, where the monetary flow from the banking sector gives rise to an equivalent monetary reflux through the consumption expenditures of households, unless households choose to hold savings in the most liquid form of cash or bank deposits, in which case the latter holdings can short-circuit the process.

Let us now consider a world of financial deepening represented in Figure 5.3. Households can now choose a whole portfolio of financial assets issued directly by firms, such as corporate stocks or bonds, or other forms of financial instruments offered by investment banks, or even through the intervention of non-bank financial intermediaries. These liabilities may not be considered good substitutes for commercial bank deposits and thus are not normally acceptable as means of payments. This broad spectrum of institutions constituting the financial markets is depicted in Figure 5.3 as a separate space into which household saving flows out of household income and these institutions do engage in financial intermediation. However, this intermediation is not between savers and investors, as it is normally described by the mainstream theory of loanable funds, because investment, or the new production and accumulation of capital goods, has already occurred. After production has taken place, what the financial markets do during the reflux phase is to bring together households which have chosen to save in the form of less liquid assets, and firms in search of long-term or 'final' finance, in order to allow the latter to extinguish their 'initial' short-term debts vis-à-

Figure 5.3 The heterodox conception of banks as creators of credit money with organized financial markets

vis the banking sector. This is described by the arrows going from households choosing to save a portion of their incomes being channelled into the financial markets, and firms simultaneously accessing these savings for final finance.

There are numerous complications to this heterodox macroeconomic analysis of banks as creators of credit-money that can be added to offer a greater degree of realism. For instance, we can include in the monetary reflux not only the principal of the loan made out to firms that must be reimbursed, but also an analysis of the interest rate spreads from which banks traditionally make profit. In addition, one can also consider the case where households enter into debt vis-à-vis the banking sector to obtain consumer loans or take out mortgages. Also, from the portrait just described of the relation between banks and firms, on the one hand, and firms and households and the financial markets on the other hand, we have excluded the role of the government and the central bank. However, these are complications that can and have been analysed by heterodox economists. In order to provide the reader with a more comprehensive understanding of the mechanics of this fundamental relation between banks and the private sector, by considering the case of a consumer loan, but also between commercial banks and the central bank and payments system, let us now describe briefly this process from the standpoint of banks' balance sheets.

Understanding the heterodox approach to banks and the modern payments system from a simple balance sheet perspective

Most mainstream textbooks treat money just as they would treat commodities: they assume that money should be scarce for it to keep its value, as if money were akin to gold, and they assume that it is the role of the central bank to make sure that this is so. If there is too much money or if the stock of money grows too fast, according to the mainstream story, there will be an increase in the general price level and the exchange rate of the domestic currency relative to other currencies will depreciate.

We have seen that mainstream economists assume the existence of a money creation mechanism based on the so-called money multiplier process and fractional-reserve banking system. They assume that in order to be able to grant loans and issue money, commercial banks must first acquire a special kind of money – reserves at the central bank – because financial regulations in a number of countries require banks to hold a certain fraction of the money deposits of their customers in the form of a specific kind of assets, namely,

reserves at the central bank. According to this story, deposits of agents at banks are thus a multiple of these reserves, roughly speaking the inverse of this required percentage; say, a multiplier of 20, if the required percentage of reserves to bank deposits is 5 per cent. By controlling the amount of reserves that commercial banks have access to, it is said that the central bank has the ability to control the supply of money in the whole economy.

The causality of this mainstream story is thus the following: depending on its objectives, mainly concerned with a stable aggregate price level, the central bank creates a certain amount of reserves. This then allows the creation of a multiple amount of loans and deposits, the latter with the addition of banknotes, constituting the supply of money, which hopefully is in line with the needs of the economy.

There are however countries where there are no reserve requirements; this obviously makes the mainstream story rather questionable if not meaningless. The creation of money must follow some other mechanism. It is the purpose of the following subsections to examine this more realistic mechanism.

Transactions of the private sector

As we have seen from our analysis of the monetary circuit, the process of money creation is simple yet fascinating. Money creation – the creation of bank deposits – relies on three key elements: the willingness of banks to take risks and grant loans, the creditworthiness of borrowers, and the willingness of borrowers to go into debt and take a loan. No more is required. Banks do not need to hold gold and neither do they need to hold reserves at the central bank. Money is created *ex nihilo*.

Suppose that an individual wishes to buy a new car worth $30 000 and needs to borrow to do so. This person will need to show that they are creditworthy, for instance, by showing that they have a regular income, that this income is likely to be large enough to make the monthly payments, and that interest and principal have been paid on previous loans (that is, the person has a good credit record).

What happens next? As pointed out in our previous discussion, the loan is created at the stroke of a pen, or rather by punching a couple of keys on the banker's computer. As the bank grants the loan, there is a simultaneous creation of a bank deposit: money gets created. This is shown in Table 5.1, which shows the changes in the balance sheet (the T-account) of the bank of the borrower: the bank now has $30 000 more in loans on the assets side

Table 5.1 Changes in the balance sheet of a bank that grants a new loan

Bank of the car purchaser	
Assets	*Liabilities*
Loan to car purchaser +$30 000	Deposit of car purchaser +$30 000

Table 5.2 Changes in bank balance sheets after the payment is made

Bank of car purchaser		*Bank of car dealer*	
Assets	*Liabilities*	*Assets*	*Liabilities*
Loan to car purchaser +$30 000	Debit position at clearinghouse +$30 000	Credit position at clearinghouse +$30 000	Deposit of car dealer +$30 000

of its balance sheet, but simultaneously, on the liabilities side, there is an increase of $30 000 in bank deposits. The car purchaser now has a bank debt of $30 000, which from the standpoint of the bank is an asset, but at the same time the bank now owes a $30 000 deposit to the car purchaser. This is why it is on the liabilities side of the bank's balance sheet.

What then happens next? The individual obtained a car loan because they wanted to buy a car. So the purchaser goes to the car dealer, most likely with a certified cheque, and once all papers are signed, drives off with the car, while the car dealer rushes to deposit the cheque in their bank account. Once this is done and the cheque goes through the payments system, the new balance sheets are shown in Table 5.2.

All banks are members of a payments and settlement system, either directly, or indirectly in the case of small banks that use the account of a larger bank. Indeed, their participation in the payments system is one of the key services rendered by banks. As payments go through banks, bank deposits move from one account to another. These payments are centralized at a clearinghouse, which keeps the tabs, so to speak. As the cheque (in paper form or electronic form) clears the payments system and goes through the clearinghouse, the $30 000 are taken away from the car purchaser and end up in the bank account of the car dealer. However, now the bank of the car purchaser owes $30 000 dollars to the clearinghouse, while the clearinghouse owes $30 000 to the bank of the car dealer, which is what Table 5.2 illustrates.

Table 5.3 Changes in bank balance sheets when banks lend to each other

Bank of car purchaser		*Bank of car dealer*	
Assets	*Liabilities*	*Assets*	*Liabilities*
Loan to car purchaser +$30 000	Loan taken from bank of car retailer +$30 000	Loan to bank of car purchaser +$30 000	Deposit of car dealer +$30 000

In all payments systems, amounts due between participating financial institutions must be settled at least by the end of the day. How this will be done depends on the institutional set-up, which is specific to each country. In the simplest case, the bank that is in a credit position at the clearinghouse – the bank of the car dealer – will grant what is called an overnight loan (at the overnight rate of interest) to the bank that is in a debit position at the clearinghouse – the bank of the car purchaser. This is the overnight market, also called the interbank market, since it involves banks and a few large financial institutions. Banks will lend to each other as long as participants to the payments system have confidence in each other. Table 5.3 illustrates this situation.

As was the case with the loan to an individual, we see that the banking system relies on trust and creditworthiness. Banks must have sufficient confidence in other banks. When banks start lacking trust, the overnight market so described, where banks in a daily surplus position at the clearinghouse lend funds to banks that are in a negative position, will freeze and banks will decline to lend to each other. This happened in Europe in August 2007, when all financial institutions were scared to make overnight loans to German banks, because of the failure of two German banks. Fears spread to the rest of the world and overnight markets lost their fluidity elsewhere as well, as banks became reluctant to lend large amounts to each other.

What then happens if the overnight market does not function properly or if, for some reason, a bank in a negative position at the clearinghouse cannot get an overnight loan from some other bank? Does the payment made to the car dealer still go through? It will, and this is where the central bank plays its role of lender of last resort. In this case, using again the two banks described in Tables 5.1 and 5.2, the central bank makes an overnight loan to the bank of the car purchaser, thus allowing it to settle its position at the clearinghouse, as shown in Table 5.4. And what happens to the bank of the car dealer? If it declines to lend its surpluses at the clearinghouse, it has no other choice than to deposit its surpluses in its account at the central bank. The deposits of

Table 5.4 Changes in the balance sheet of banks and the central bank when banks decline to lend to each other

Bank of car purchaser		*Bank of car dealer*	
Assets	*Liabilities*	*Assets*	*Liabilities*
Loan to car purchaser +$30 000	Loan taken from the central bank +$30 000	Deposit at the central bank +$30 000	Deposit of car dealer +$30 000

Central bank	
Assets	*Liabilities*
Advance to the bank of car purchaser +$30 000	Deposit of the bank of car dealer +$30 000

the bank of the dealer at the central bank are what mainstream authors call reserves; central bankers now refer instead to clearing balances or settlement balances. Table 5.4 also illustrates the fact that the size of the balance sheet of the central bank will balloon any time overnight markets do not function properly.

As noted previously, there are many possible set-ups for payments and settlement systems. In the set-up assumed so far, unless the overnight market collapses, all the activity occurs in the clearinghouse, which can be run by a private entity, owned by the bankers' association for instance. Another possible set-up, often assumed in textbooks and actually existing in several countries, is that clearing and settlement occurs on the books of the central bank. In that case, payments can only go through, and hence settlement occurs, if the bank making the payment – here the bank of the car purchaser – already has deposits at the central bank (if it has reserves). Table 5.5 illustrates this situation: the bank of the car purchaser sees its reserves at the central bank diminished by $30 000, while those of the bank of the car dealer get augmented by the same amount.

When an individual bank starts to run out of reserves, it will have to borrow funds on the overnight market, thus borrowing the funds from banks that have a surplus of reserves (this is the federal funds market in the United States); or it might borrow the reserves from the central bank, as illustrated in Table 5.4. Thus the role of the central bank is not to put limits on the creation of reserves; its role is purely defensive: it needs to provide enough

Table 5.5 Changes in balance sheets when the central bank acts as the clearinghouse

Bank of car purchaser		*Bank of car dealer*	
Assets	*Liabilities*	*Assets*	*Liabilities*
Loan to car purchaser +$30 000 Reserves at central bank –$30 000		Reserves at the central bank +$30 000	Deposit of car dealer +$30 000

Central bank	
Assets	*Liabilities*
	Deposit (reserves) of the bank of car dealer +$30 000 Deposit (reserves) of the bank of car purchaser –$30 000

reserves to ensure that the payments system runs smoothly. In countries where the clearing occurs essentially through the clearinghouse, and where payments are netted out at the end of the day, there is no need for reserves, and hence, unless some regulation imposes required reserves as a fraction of some measure of assets or liabilities, there will be zero reserves, as is the case of Canada, for instance. By contrast, in countries where the clearing occurs through the central bank and where payments are settled in real time as they go through, banks will have to hold reserves at the central bank, which will then act as the clearinghouse, ensuring that there are enough clearing or settlement balances to absorb the fluctuations in incoming and outgoing payments.

Transactions of the public sector

What should be noted is that the total amount of clearing balances (or reserves) in the banking system is a given as long as all transactions occur between private agents. As is obvious from Table 5.5, any increase in the reserves of one bank will be compensated by the decrease in the deposits of another bank. Thus the overall amount of reserves in the banking system can only be changed if a transaction occurs with the public sector, that is, when the central bank is involved in one of the transactions. Such a transaction is in fact already described by Table 5.4. In this case, the central bank provides an advance to the banking sector, thus generating the creation of an equivalent amount of reserves for the banking system. More generally, if the central bank feels that there is a higher demand for reserves by the banking sector,

Table 5.6 An open-market operation when the central bank wishes to increase the amount of reserves or to compensate for a previous fall in reserves

Commercial bank		*Central bank*	
Assets	*Liabilities*	*Assets*	*Liabilities*
Treasury bills –$10 000 Reserves at central bank +$10 000		Treasury bills +$10 000	Deposit of commercial bank +$10 000

additional reserves can be created by making advances to some banks, for one night, one week, one month, or perhaps even three years as was done by the European Central Bank at the height of the euro area crisis in 2011.

However, there are other ways in which reserves are created (or destroyed). As reflected in mainstream textbooks, reserves are created whenever the central bank purchases assets from the private sector. Table 5.6 provides such an example. It is assumed that the central bank purchases a government security from the banking sector; for instance, a Treasury bill that had been issued earlier by the government and bought by a bank. This transaction, whether it is outright or whether the central bank promises to sell it back within a period of time (in which case it is a repurchase agreement, a repo), leads to the creation of new reserves for the banking sector. This type of transaction is called an open-market operation. Of course, it can go the other way; for instance, when the central bank sells the Treasury bills that it holds to the private sector, in which case reserves are destroyed.

The central bank is usually the fiscal agent of the government. This means that the central bank is empowered with the responsibility of selling the securities that the government issues when it borrows funds; it also means that the central bank manages the cash balances of the government and, in particular, it implies that the government has an account at the central bank. Consequently, any time there is an outgoing or incoming payment involving the government deposit account at the central bank, there will be a creation or a destruction of reserves (Box 5.1). Central bankers call these the 'autonomous factors' that affect the amount of reserves in the banking system.

Take the example of a civil servant receiving their monthly pay, assuming that it comes out of the account of the government at the central bank. Table 5.7 illustrates this case: the bank account of the civil servant will now increase by $5000. As the payment goes through the clearinghouse, and is settled, the

Box 5.1

WHERE DO GOVERNMENT DEPOSITS COME FROM?

In Table 5.7, we assumed that the government had deposits at the central bank, thus allowing it to pay its civil servants. However, where do these deposits come from? Several answers are possible. First, the central bank could make a loan to the government, just as a bank provides loans to producing firms, but this is universally forbidden today. Second, the government could issue securities, purchased entirely or in part by the central bank, with the government thus acquiring deposits at the central bank; besides Canada, few countries, however, proceed in this manner. A third possibility is for the government to issue securities, purchased (on the primary market) by banks or bond dealers: this is the standard procedure. The proceeds of the sale are then brought back to the deposit account of the government at the central bank. In this case, banks wind up with a negative position at the clearinghouse or with a loss of reserves; the central bank must either provide advances to the banking sector or purchase back the securities on secondary markets, if it wants to keep the overnight interest rate at its target level. It would then have been simpler for the central bank to go with the second possibility and buy the securities right away (on the primary market). A fourth possibility is when taxes are being paid by households and firms, as these payments feed the deposit account of the government at the central bank. It should be noted, however, that if the incoming taxes are greater than government expenditures, the banking sector overall will again be in a negative position at the clearinghouse, and will need to be provided with advances by the central bank.

Table 5.7 Changes in balance sheets induced by government expenditure

Commercial bank		*Central bank*	
Assets	*Liabilities*	*Assets*	*Liabilities*
Reserves at central bank +$5000	Deposit of civil servant +$5000		Deposit of government –$5000 Deposit of commercial bank +$5000

government deposits at the central bank decline by $5000, while the reserves of the bank of the civil servant increase by $5000. Thus, when the government makes a payment to the private sector, through its account at the central bank, this creates reserves. Things go in reverse gear if the civil servant has to pay their income taxes; say, at the rate of 40 per cent. The deposits of the civil servant will fall by $2000 and, if the proceeds are deposited in the account of the government at the central bank, the reserves of the banking system will be reduced by $2000 (Wray, 2012, Ch. 3).

Box 5.2

ARE THERE LIMITS TO CREDIT CREATION?

If bank reserves (the amount of their deposits at the central bank) are not a constraint on money and credit creation, as argued in the chapter, is there another supply constraint? A number of economists, orthodox and heterodox alike, believe that the equity of a bank (its own funds: the capital of its owners) provides such a constraint. This is based on the notion that regulations devised by the Bank for International Settlements (BIS) are such that bank loans can only be a certain multiple of bank equity. When a loan defaults, the bank equity is reduced by the size of the defaulting loss. Large losses on previous loans may thus cause trouble, either because the entire equity is wiped out (in which case the bank becomes insolvent and would need to be closed down) or simply because the remaining loan book value exceeds the allowed multiple over bank equity. In the latter case, the bank will be restrained in making new loans. However, in general, a bank which is on the edge of the BIS regulation can always make more loans if it wishes to do so: it suffices to accumulate a portion of its profits in the form of retained earnings or to sell new shares, to increase the size of its equity.

The lesson to be drawn here is that commercial banks are the institutions that grant loans and create money *ex nihilo*. There is no constraint on how much can be created, with one exception. A banker must keep the trust and confidence of depositors and of fellow bankers, and so must make sure that the number of 'non-performing loans' – loans on which borrowers default, thus creating losses for the bank – is minimized, to avoid the arising of suspicion. Commercial banks do not need central bank reserves to grant loans. On the contrary, the role of the central bank is to make sure that there is the right amount of reserves in the banking system, to ensure that the payments system is running smoothly. The central bank will react to changes in the 'autonomous factors' affecting reserves, which we discussed above either by pursuing open-market operations or by providing advances to the banking sector. To sum up, we may say that the supply of money is endogenous, responding to the demand of the economy, and that the supply of reserves is also endogenous, responding to the needs of the payments system (Box 5.2).

Concluding remarks

The heterodox view of money and banking stands on its head much of the mainstream theorizing on the role of money and banks that still reign supreme in popular textbooks. Money is not a commodity that is dropped

exogenously from a helicopter to render more efficient the exchange of commodities and services. Money is a means of payment that permeates social relations and permits economic agents in a community to free themselves from the constraints of scarcity that a barter economy imposes. Money is not some predetermined object but, in its essence, an endogenous creation that is issued by institutions established or licensed by the state. This money emerges through a balance sheet operation associated with changes in a third-party liability, whether it is the state through its central bank liability, or through the strategic role played by banks via credit-money creation reflected in changes in their deposit liabilities. As long as there are creditworthy borrowers, credit-money creation is demand-driven. It can never be supply-constrained as claimed by the mainstream. Moreover, because banks are not reserve-constrained as a group and money creation is an outcome of the interaction between an individual borrower and a bank, money cannot be an exogenous variable as normally depicted in the textbooks. Indeed, bank credit follows a circular process of money creation and money destruction and, therefore, in contrast to the mainstream, which emphasizes its role in facilitating exchange, the most crucial social feature of money is associated with its financing of production, where commercial banks have played a central role since the nineteenth century.

NOTE

1 The 'real bills' doctrine rested on the belief that prudent banking practices required that commercial banks engage only in short-term lending, in the sense that banks would discount commercial bills or promissory notes on the basis of collateral representing 'real' goods engaged in the production process. This essentially meant that commercial banks should passively accommodate the 'needs of trade' by financing the short-term circulating capital requirement, such as the wage bill, and not finance the purchases of fixed capital assets. For further discussion, see Humphrey (1982).

REFERENCES

Graziani, A. (2003), *The Monetary Theory of Production*, Cambridge, UK: Cambridge University Press.

Humphrey, T.M. (1982), 'The real bills doctrine', *Federal Reserve Bank of Richmond Economic Review*, **68** (5), 3–13.

Jakab, Z. and M. Kumhof (2015), 'Banks are not intermediaries of loanable funds – and why this matters', Bank of England Working Paper, No. 529.

Parguez, A. (1975), *Monnaie et macroéconomie: théorie de la monnaie en déséquilibre*, Paris: Economica.

Parguez, A. and M. Seccareccia (2000), 'The credit theory of money: the monetary circuit approach', in J. Smithin (ed.), *What is Money?*, London, UK and New York, USA: Routledge, pp. 101–23.

Peacock, M. (2013), *Introducing Money*, London, UK and New York, USA: Routledge.

Realfonzo, R. (1998), *Money and Banking: Theory and Debate (1900–1940)*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Rochon, L.-P. and M. Seccareccia (eds) (2013), *Monetary Economies of Production: Banking and Financial Circuits and the Role of the State*, Cheltenham, UK, and Northampton, MA, USA: Edward Elgar Publishing.

Smith, A. (1776/1937), *An Inquiry into the Nature and Causes of the Wealth of Nations*, New York: Modern Library.

Wray, L.R. (2012), *Modern Money Theory: A Primer on Macroeconomics for Sovereign Monetary Systems*, Basingstoke, UK and New York, USA: Palgrave Macmillan.

A PORTRAIT OF ALAIN PARGUEZ (1940–)

Alain Parguez is Professor Emeritus at the Université de Franche-Comté, Besançon, France. Together with the late Augusto Graziani, Parguez has been a leading figure of what has been sometimes described historically as the Franco-Italian Circuit School. His ideas are deeply rooted in the works of French Keynesian economists of the post-war period, the most noteworthy being Jean de Largentaye and Alain Barrère. However, his writings find inspiration in the macroeconomic and monetary views of Karl Marx, John Maynard Keynes, Michał Kalecki and Joan Robinson, as well as the monetary ideas of such heterodox writings going back to the Banking School in the nineteenth century.

His ideas on money and the role of banks in the monetary circuit were first espoused in an important book titled *Monnaie et macroéconomie: théorie de la monnaie en déséquilibre* (1975). These ideas on the monetary circuit were further developed in a series of publications that followed his book, particularly in the two French journals *Économie appliquée* and *Économies et sociétés*, where, in the case of the latter, he had edited a special series *Monnaie et Production* that lasted from 1984 to 1996. More recently, a book came out in his honour titled *Monetary Economies of Production: Banking and Financial Circuits and the Role of the State* (Rochon and Seccareccia, 2013), which celebrates his contributions to political economy.

According to Alain Parguez, the mainstream views on money and banking lead to a world of institutionalized scarcity, because they start from the misleading premise that money should be a scarce commodity and that banks are mere intermediaries between savers and investors. In contrast to the mainstream austerity perspective, Parguez starts from the view that there are no supply-side limits to the creation of money, because the latter can be created *ex nihilo*. Therefore, the building of models of modern economies where banks are mere intermediaries whose loan advances depend either on past savings or on the amount of base money created by the central bank, rather than being conceived as creators and destroyers of money, can lead to catastrophic policy consequences. An example of this misguided policy, based on such erroneous theories of commercial banking, is the attempt to kick-start economic growth since the 2007–08 financial crisis through the policy of quantitative easing. This led to an explosion of reserves in the banking system, but without significant effects on economic growth, with the possible exception of sustaining asset prices by keeping the level of interest rates at their lower bound.

In a similar fashion, Parguez has been a staunch critic of the institutional structure of the euro area because of the formal severing of any direct link between monetary creation through a supranational central bank (the European Central Bank) and the actions of the national fiscal authorities. He has argued that the original architects of the euro tried to create a monetary system not unlike that founded on gold, going back to the nineteenth century, with results that have been just as catastrophic. Although Alain Parguez has been an ardent supporter and advocate of the monetary circuit approach, over the years he has done much to promote exchanges with other heterodox theorists of money within the broad post-Keynesian tradition, for instance with economists such as Paul Davidson and Basil Moore, as well as with writers associated with Modern Monetary Theory (MMT), such as Warren Mosler and L. Randall Wray.

EXAM QUESTIONS

True or false questions

1. The heterodox theory of money rejects the view that money should be considered merely a special type of commodity invented for the primary purpose of facilitating exchange. Money is the result of a balance sheet operation resulting from the appearance of a third-party liability of either a private commercial bank or a central bank.
2. Money creation arises when the central bank prints more money from which a multiple expansion of deposits can take place within the banking system. Hence, money is supply-determined and the only limit to the money supply expansion is the amount of reserves held by the banks and the amount of cash held by the public.
3. In their essential role, banks are storehouses of money or mere financial intermediaries specialized in deposit taking upon which they can then fund creditworthy borrowers desiring to invest within an organized market for loanable funds.
4. The distinction between initial finance and final finance is crucial to understanding the process of monetary circulation. The first (initial finance) describes the creation or injection of money through bank loans as this new money generates income flows, while the latter (final finance) describes the reflux side or the monetary destruction as the borrowing unit reimburses its debt.
5. Liquidity preference is the withholding of private spending in liquid form as bank deposits. The accumulation of bank deposits can 'short-circuit' the flux/reflux mechanism and prevent the business sector from reducing the initial bank debt.
6. The own funds of the bank (or its paid-up capital) should appear on the asset side of the balance sheet of a bank.
7. Tax revenues are a prerequisite to public spending.
8. Within the mainstream view, the role of the central bank is primarily to prevent the excessive expansion of the money supply by setting a floor on the holding of bank reserves. In contrast, heterodox economists emphasize the lender of last resort or accommodative function of the central bank in supplying reserves for the proper functioning of the payments system.
9. When the government spends more than it receives in tax revenues within a given period, thereby running a budget deficit, the rate of interest on the overnight market for funds goes up, in which case the central bank must remove reserves from the banking system.
10. In a modern monetary system, while a single commercial bank can be reserved-constrained, the private banking sector, as a whole, can never be, unless affected by the autonomous actions of the government in reducing the amount of overall reserves in the system.

Multiple choice questions

1. What seems truer about the origins of money?
   a) In a barter economy in which individual exchangers benefit from the division of labour, money was invented as medium to minimize transactions costs relating to the exchange of goods and services in a market context.
   b) Since money is an abstract unit of account, only in a modern advanced economy where there already exists a sophisticated division of labour could money actually have appeared.
   c) With the increasing division of labour and specialization, money emerges from the expanded use of precious metals technologically, such as gold and silver, because of these commodities' special characteristics of divisibility, portability, durability, fungibility and their scarcity.

d) Money emerged historically as a means of payment, in its capacity to extinguish debt obligations, of which the most important was tax liabilities to the state, regardless of the degree of sophistication of the division of labour in that economy.

2. Heterodox economists flip the theory of the so-called 'money multiplier' on its head, in which exogenous changes in base money trigger a multiple expansion of the money stock in an economy, because:

a) they interpret the causality in reverse to that of the traditional interpretation that is based on a supply-side view of bank lending behaviour;
b) base money is no longer composed of precious metals;
c) banks do not normally try to seek to extract greater profits when seeking to expand loans on the basis of the increased base money;
d) banks can only lend their excess reserves.

3. What happens to the bank equity if the bank suffers from a bad loan worth $10 billion?

a) The liability side shrinks by $10 billion.
b) The asset side shrinks by $10 billion.
c) Both the asset side and the liability side shrink by $10 billion.
d) Both the asset side and the liability side increase by $10 billion.

4. In the circulatory process that is connected with the process of monetary creation and destruction in a modern economy, which was described as a monetary circuit, identify which of the following actions pertain to the monetary 'reflux' associated with an eventual destruction of money:

a) the issuing of a bank loan to pay wages and salaries by a firm;
b) the regular payment of interest by households on a bank loan;
c) the weekly grocery purchases by a household;
d) the purchase of an already existing corporate bond by an individual through a brokerage firm.

5. The dictum that 'loans make deposits' can be described as representing the normal functioning of the banking system, because banks are not reserved constrained collectively. However, we have suggested at least one exception when 'deposits make loans'. What is this exception?

a) This exception arises during times of crisis when the public is not sufficiently creditworthy and, therefore, risk-averse bankers will make loans only if they have first secured an equivalent amount of deposits.
b) During times of strong liquidity preference reflected in the build-up of an inordinate amount of liquid deposits in the banking system, this could force commercial banks into an uncomfortable position of extending loans to prevent business insolvencies and to cut bank losses.
c) In reality, the dictum is wrong, since banks would not normally want to leverage themselves by issuing loans without the backing of deposits. The saying that 'loans makes deposits' is actually the exception and not the norm.
d) The exception arises when banks lose trust and cease to lend excess reserves among themselves in the overnight interbank market for funds.

6. What happens to the balance sheet of the bank and that of the central bank when the government issues new securities that are purchased by banks, with the proceeds of the sale being kept by the government as deposits in the banks?

a) Banks have a bigger balance sheet and so does the central bank.
b) Banks have more bonds on their liability side.
c) Banks have more bonds on their asset side.
d) Banks have more reserves at the central bank.

7. As happened during the global financial crisis, many commercial banks facing a negative position in the clearing system were unable to obtain an overnight loan from other banks which found themselves in a surplus position because of the fear of bank insolvencies. What could banks mostly likely do if they find themselves in a negative position?
   a) They could pretend that nothing has happened and announce that their books are balanced in the hope that the bank supervisory authorities do not recognize their balance sheet problem.
   b) They could sell some of their assets to the public.
   c) The could declare bankruptcy and merge with banks in a positive settlement position.
   d) They could borrow directly at the discount window of the central bank.
8. It has been shown that, while the total reserves within the banking sector remain unaffected by the volume of transactions within the private sector, the overall amount of reserves in the banking system adjusts only when transactions are with the public sector. In particular, total bank reserves go up:
   a) when the central government cuts various taxes and service charges, without cutting proportionally its overall spending;
   b) when the central bank buys government securities in the financial markets as occurred under quantitative easing (QE) interventions;
   c) when the government transfers some of its deposits from its account at the central bank to a commercial bank;
   d) if any of the above occur.
9. What happens to the balance sheets of the banks and of the central bank when households pay their income taxes by drawing on their bank deposits?
   a) Deposits of households at banks go up, but the deposits of banks at the central bank go down.
   b) Deposits of households at banks go down, but the deposits of banks at the central bank go up.
   c) Both the deposits of households at banks and those of banks at the central bank go down.
   d) Both the deposits of households at banks and those of banks at the central bank go up.
10. Following the global financial crisis of 2007–08, a number of central banks implemented a policy of quantitative easing (QE) to encourage bank lending and re-establish private spending by flooding the banking sector with excess reserves.
   a) The effect was to reduce interest rates dramatically and stabilize asset prices, thereby leading to high economic growth.
   b) The effect was to bring down the interest rate on the overnight interbank funds market to its lowest possible level, but its effect on bank lending and economic growth was marginal.
   c) The massive rise in bank reserves directly stimulated bank lending and, therefore, prevented a drop in private spending.
   d) Since fiscal policy was seen as an ineffective tool and counterproductive because it would push interest rates up, policy-makers had no other choice but to opt for QE to stimulate private spending.

6

The financial system

Jan Toporowski

OVERVIEW

This chapter explains that the financial system:

- is that part of the economy that meets the financing needs of the rest of the economy;
- is very large in key capitalist countries such as the United States and United Kingdom;
- is unstable and prone to crisis;
- in mainstream economics is confused with 'savings';
- emerges into the modern world with the financing needs of the state, subsequently with the financing needs of infrastructure and industry;
- shows up the defects of modern capitalism.

KEYWORDS

- **Balance sheet:** A financial statement showing the assets (income-generating wealth or claims on other government and other economic units) and liabilities (commitments to pay in the future).
- **Bond:** A contract recording the loan of a certain amount of money to the issuer of the bond. The bond usually specifies the interest payable on the loan, the dates when it is due, and the date when repayment is made. The interest and repayment terms are what give value to a bond. Bonds may be bought and sold on a stock exchange, allowing owners of a bond to benefit from a capital gain (or loss) if the bond is sold at a higher (or lower) price than the one originally paid for the bond. Chasing such capital gains is the art of speculation.
- **Common stock:** See **Shares**.

- **Debt:** An obligation to pay money to another person, government, bank, financial institution or firm. Debts are usually eliminated by paying (or settling them) with money.
- **Financial wealth:** A claim on another person, government, bank, financial institution or firm, requiring them to pay money. Financial wealth is usually distinguished from other forms of wealth: personal wealth (such as jewellery), land or real estate, and productive capital (factories, trucks, and so on).
- **Interest:** A payment for the use of money during a certain period. Interest is usually calculated over the period of a year. The contract for the use of that money can be a loan, bond or a short-term bill.
- **Shares or common stock:** Titles of ownership of a company. Ownership of a share in, or common stock of, a company entitles the owner of a share to a dividend or payment out of the profits of a company, and to a share of the proceeds of the company, after payments of its debts, if the company is sold off to new owners.
- **Stock exchange:** A market where bonds and shares are bought and sold.

Why are these topics important?

Finance is a part of the economy in which the financing needs of the economy are met through credit granting and the provision of financial instruments such as bonds, shares and financial futures; in other words, it is the set of institutions and practices that contain the financial wealth and obligations of the state, households and firms. That set includes banks, investment institutions and the credit operations across capital markets.

In most developed capitalist economies, the financial sector is very large. Indeed, in the case of the United States or the United Kingdom, for instance, its value and turnover are some four or five times the value of the annual national income; in other words, it is four to five times the value of everything that is produced in one year. For many economists, this is a kind of excess, which arises because those who are not bankers and financiers are all becoming 'financialized' and have their incomes burdened with the payments on this growing debt from national income (see Chapter 18). So, for example, if total debt is four times the national income, then roughly four times the rate of interest has to be taken out of national income to pay interest on that debt. Yet, that is an incorrect way of viewing this, since most debts are claims on other debts. In other words, the money that a bank pays on its debts comes largely from the money that its customers pay on their debts toward the bank. The net debt (that is, after cancelling out debts backed by other debts) is therefore negligible. So what is actually taken out of national income to service debt is proportionate to net debts rather than the total of all debts.

Box 6.1

HEDGING YOUR ASSETS, THE KEYNESIAN WAY

In *The General Theory of Employment, Interest and Money*, John Maynard Keynes (1936) put forward the idea that the most efficient 'hedge' or security against changes in the value of any asset is money (cash or bank deposits). This is because the money value of cash or a bank deposit will always be the same, whereas a longer-term bond or a share can fluctuate in value. If I hold a portfolio of shares that fluctuates in value between $100 and $150, then the range over which the portfolio value changes is something like 50 per cent. If, however, I hold half of that portfolio in the form of cash, then with the same fluctuations in the share prices, my portfolio will only fluctuate between $100 and $125, and the range of variation is now only 25 per cent. Not only is the value of my portfolio more stable, but I also now have a money reserve that allows me to pay unexpected expenses without having to sell off shares, possibly at a time when their price may be low. Keynes thought that as bond or share prices rise, holders of financial assets would prefer to hold more cash, because they would be unwilling to buy shares at high prices. This he called the 'speculative demand for money'. The holding of money against unexpected expenses he called the 'precautionary demand for money'. But both of them presuppose the existence of a sophisticated financial system in which wealthy people hold financial assets.

Most of this size and remarkable activity of the financial system are due to two circumstances. The first is the various processes of 'hedging', which occurs when individuals take financial positions (that is, enter into financial contracts) to protect themselves against changes in the value of the assets, or more specifically to offset possible losses or gains, essentially a way in which bankers and financiers make money from each other (Box 6.1). At times of economic hardship, the sums involved, and the incomes derived from such contracts, expose such individuals to accusations of unfair profiteering from the misery of those affected by that hardship.

The second cause arises through the emergence of 'secondary markets' for longer-term financial assets. Holders of stocks and shares rarely buy a bond, and hold it until it is repaid by the issuer of the bond, in particular if the bond has 20 or 30 years left before repayment. Shares in any case are not formally repaid unless a company is taken over and its shares bought for cash. 'Secondary markets' for financial securities exist in which bondholders and shareholders can buy and sell stocks and shares among themselves. Such markets also allow holders of bonds or shares to use such financial assets as security for bank loans, because a bank can immediately verify the current market value of stocks in order to decide how much to lend

against such security. Holders of bonds or shares therefore have easy access to borrowing. But the borrowing is usually for the purpose of buying other financial assets so that the cost of the borrowing comes from other financial income, rather than being a direct drain on national income. In this way, gross debt or interest payments exaggerate the burden of financing in an economy.

Added to this complexity is the growing prominence and international integration of financial markets: capital and banking markets in the United States and the United Kingdom now also serve to transfer capital between other countries of the world, so-called intermediation: money-capital from Russia, for example, appearing in New York on its way to Latin America or China; or American pension funds holding bonds issued by American corporations in offshore financial centres, such as the Bahamas, London or Ireland. This growing international intermediation makes it appear as though the average Briton or American is supporting a huge amount of debt, when in fact much of it is being serviced by many more people in other countries of the world.

Financial markets have come to the fore in economic discussions since the financial crisis in the United States in 2007–10, and owing to the distinctive role that banking and financial markets played in the European economic crisis since 2010. The collapse of Lehman Brothers in September 2008, and the failure of Northern Rock and the Royal Bank of Scotland in the United Kingdom, were followed by the deepest economic recession since the Second World War (Box 6.2). The succession of events looked remarkably like the Great Crash of 1929, which gave rise to the Great Depression of the 1930s. To be clear, the situation of finance in the twenty-first century's first international economic crisis is complex, with different causal factors at work in different countries: while the American crisis was characterized by a failure of private sector debt, in Europe government debt has been in crisis. This chapter aims to explain some of that complexity.

More generally, a closer examination of economic history over the last 200 years shows us that capitalist economies are unstable and that debt problems and banking crises are common at the very moment economic activity begins to contract, not to mention when a crisis occurs. This makes an understanding of finance crucial for understanding macroeconomics, or the way in which economies as a whole work.

Box 6.2

THE 2008 FINANCIAL CRISIS

On 15 September 2008 the American investment bank Lehman Brothers filed in the United States (US) courts for protection from its creditors. This was rapidly followed by the collapse of prices in the housing market and economic recession in the US, from which the economy has struggled to recover. A couple of years later the crisis started in Greece, and Europe too succumbed to depression. The parallels with the 1929 crash, which was followed by the Great Depression of the 1930s, are immediately obvious. The narrative runs as follows: with the deregulation of banking in the 1990s, bankers had colluded to create collateralized debt obligations (CDOs) that no one understood, based on indebting poor people in the housing market; and when those poor people could no longer pay, they were dispossessed of their homes, and the CDO market collapsed, leading to bank failures. Financial crisis has come to be associated with economic depression. Even mainstream economists have questioned why no one thought of looking at the financial system as a source of economic disturbance (or 'shocks'), while heterodox economists have pointed the finger of blame at economic theories that do not incorporate finance and have rushed to declare themselves in the forefront of those who had anticipated the financial crisis.

As with most things in economics, the situation was and remains a bit more complex than this summary suggests. The housing market in the US and in the United Kingdom for decades has been and remains a place of suffering and misery for poor people. But most of the collateralized home loans were for middle-class housing that remains largely unaffected by the crisis. What came to be called the Global Financial Crisis was anything but global, since most people in the world, in China, India and the developing countries, were unaffected by the crisis. And the crisis in the European Monetary Union (EMU) had nothing to do with the housing market in those countries, and everything to do with the hostile environment for government finance created by the EMU. The narratives in mainstream and heterodox economics overlook the key transmission of financial crisis through corporate finance. The capital markets for long-term finance (stocks and shares) in the US and the United Kingdom had been in difficulties since the 1990s, as the demand for long-term bonds dried up because pension funds and insurance companies were no longer buying them on any significant scale. The 'dot.com bubble burst' in 2001 was an early indication of these difficulties.

When the US central bank, the Federal Reserve, responded by lowering interest rates and easing borrowing conditions, US and multinational corporations borrowed heavily in the short-term money markets to finance merger and acquisition activity. The strategy was to refinance (or fund) the borrowing after the corporate takeover with the issue of long-term bonds or shares. However, banks started having problems selling the CDOs that they were using to fund their lending. When suspicions arose in the money markets that borrowers were in difficulties, lending in those markets froze up. The corporations found themselves with large short-term debts that they could neither roll over in the money markets ➡

⬅

(borrowing to repay the earlier debt) nor refinance in the capital market. Faced with the requirement to repay their short-term borrowing, the corporations took the obvious step to maintain payments. They cut their investments. It is the reduction in private sector investment in productive capacity, rather than any fall-off in household consumption, that pushed the US and then Europe into recession.

The US authorities responded quickly with the Troubled Asset Relief Program (TARP) in October 2008. This bought over US$400 billion of CDOs and other bonds from US banks. Together with subsequent quantitative easing programmes and currency swap agreements between central banks, the TARP relieved the situation for banks. However, business investment remains low and is the cause of the economic stagnation that hangs over Europe and North America.

The traditional mainstream view

Mainstream economic theory gives a very poor account of the operations of the financial sector. Several reasons can explain this neglect, of which five are discussed here.

First, finance appears in textbooks principally as portfolio theory, that is, the theory of 'rational' or 'optimal' portfolio choice, meaning the choice households make about how to allocate their wealth, between, say, bonds or another asset. This choice is supposed to be determined by the returns from holding such an asset. Rational portfolio choice involves getting the highest possible returns from the portfolio, given the possibility of loss that may arise in all uncertain situations (and the future of stock prices is always uncertain).

Portfolio theory is supposed to be a way of calculating how to make the largest possible amount of money out of savings or financial investments. But like any system that promises to make you rich overnight, portfolio theory has all the scientific rigour and insight of seventeenth-century alchemy that promised to turn base metals into gold. The theory has no substantial explanation of how returns are generated in the economy, or the relationship of finance and credit to the rest of the economy. Portfolio theory can therefore be safely disregarded. The only reader who may need it is one who will be studying finance for specialist courses. Even then, such a reader, who might eventually work in the financial sector, will find portfolio theory virtually useless for guiding actual financial operations. Such operations are determined by the financial strategies of financial institutions, rather than by individual investors' portfolio choices.

Second, the neglect of the importance of finance by mainstream economics is largely due to the success of what is called the Modigliani–Miller theorem in finance and economics, according to which the value of a company is not affected by its financing structure. In other words, the theorem, developed by Modigliani and Miller (1958), claims that the value of a company is independent from the proportion of debt or equity used to finance the company. If this is the case, then there is no need to discuss finance. Yet, this proposition holds true only under very special circumstances or assumptions, such as perfect knowledge of the future, the absence of market 'distortions' such as taxes, and perfect liquidity. But somehow these special circumstances got lost in the process of the acceptance of this theorem into the broader economic curriculum, with the result that the importance of the financing of business was simply forgotten.

Third, another way in which finance appears in mainstream theory is as household savings, resulting from mainstream economists seeking, since the 1970s, to introduce microeconomic foundations (or microfoundations) into the dominant economic theory (see the Conclusion of this volume). But finance is much more complex than simply discussing how households place their savings. A variation of mainstream economics, the New Classical doctrine supposes that the economy consists of only households who make calculated decisions about production and exchange, including the sale of their labour, and how much to save for retirement. These savings are then supposed to determine the amount of investment that the economy undertakes, since according to mainstream theory savings determine investment. But, since households control all the available resources in the economy, there cannot possibly be any unemployed resources: holding unemployed resources (including labour) would be irrational from the point of view of the household. If there are any unemployed resources, it is merely because households are looking for the next opportunity to trade.

Accordingly, there is no unemployment. And recall that a flexible price system ensures that resources are exchanged to the point where everyone has the goods they prefer. Production is undertaken using only 'real' resources (labour, materials and equipment, but not money), since this is a real exchange economy and not a monetary economy of production. Savings therefore consist of claims on future income from production, that is, entitlements to consumption goods that will be consumed during the saver's retirement.

This New Classical theory is plausible only because it plays on the vanity of the university professors who promoted it. This is how rational individuals

are supposed to behave in the conditions postulated by the theory. Any other theory must therefore assume that people are irrational, and people, especially university professors, are not irrational.

Unfortunately, this theory became widespread at a time of growing financial and banking instability: the United Kingdom experienced a major stock market crash in 1974, followed by bank failures and then the failure of savings and loan associations in the United States at the end of the 1970s. In 1982, the international banking system nearly collapsed, because many indebted developing-country governments, in Mexico, Brazil, Argentina and Poland, could not pay their debts.

How did respectable university professors of economics respond to the financial and banking instability? They did so by arguing that the instability reflected some perceived objective risk of loss that exists in the world. For instance, in a famous model that was supposed to explain how banks work, Douglas Diamond, of the University of Chicago, and Philip Dybvig, of Yale University, postulated two types of households (or agents): borrowers and lenders. This model assumed that some agents took in the savings of some households (deposits) and then lent them out to other households (loans). A potential problem arises, however, since households are supposed to be able to withdraw their savings at any time. But the savings that were lent to other households (the borrowers) were lent for longer periods of time, and subject to the possibility of loss. So if the deposits were lent out for, say, a period of five years, what happens when the depositor wants to withdraw their savings after only one year?

This is supposed to explain bank failures; that is, when banks cannot pay out savings deposited with them (see Diamond and Dybvig, 1983). Bankers, of course, know what makes them fail, so the Diamond and Dybvig (1983) model did not offer bankers any new insights. In practice, bankers also have techniques for managing liquidity, for instance holding bonds that may be sold to raise cash quickly. But the important thing about the Diamond and Dybvig (1983) model was that it showed the conditions under which household banks, run by individuals thinking like university professors of economics possessing rather limited worldly experience, could get into difficulty.

Fourth, at around the same time, Joseph Stiglitz started to publish a series of models arguing that the problems of banks were due to inadequate information on the part of those households/agents on the risk characteristics (ability to repay loans) of their customers. Since the borrowing agents are supposed to know more about their prospects of repaying their debts than

the banks, he called this problem one of 'asymmetric information' (see Stiglitz and Weiss, 1981): one party has more information than the other.

This phrase became the call-sign of Stiglitz's followers, the so-called New Keynesians. Stiglitz then used this information problem of banks to explain phenomena such as unemployment and, later on, the weak development of banking in developing countries. According to Stiglitz, banks based in the United States and the United Kingdom did not know the true economic conditions in the developing countries in which they had invested, and those countries did not have financial systems that could cope with these banks' demands to get back their money.

Asymmetric information rose to become a dominant theme of macroeconomics. For instance, by the end of the 1980s, financial and economic instability was sufficiently obvious for (the then) Princeton University Professor Ben Bernanke to get together with Mark Gertler, of New York University, to attempt to explain this instability. They created a model, relying on asymmetric information, showing how financial markets may affect the economy at large through fluctuations in the 'net worth' of borrowers, that is, the value of their assets, after taking away the value of their liabilities (see Bernanke and Gertler, 1989). The problem of this approach is that the authors made no attempt to explain the processes by which these changes in net worth arose, other than by reference to the all-pervasive asymmetric information. Net worth decreases when stock and bond prices fall. Not knowing when this will happen is supposed to be due to asymmetric information.

Now, net worth certainly fluctuates with the business cycle, but these processes are of crucial importance for the theory. Only by correctly understanding these processes is it possible to determine whether changes in net worth are symptoms of the business cycle and determined by them or, as Bernanke and Gertler (1989) argue, cause business fluctuations. The implications can be very important.

Close to the New Keynesians is the school of behavioural finance, explaining that, in states of ignorance and uncertainty, people make decisions based on rules of thumb, or emulate others who they think have superior knowledge. This can lead to economic and financial decisions that take the economy away from the efficient outcomes and equilibrium beloved by mainstream academic economists. Led by Robert Shiller, behavioural finance theorists have injected a dose of realism into the discussion about financial decision-making (see Shiller, 2000, 2003). However, they provide little institutional detail to disentangle the web of financial structures that constitutes the finan-

cial system. This is not really surprising, since their purpose is not to reconstruct the financial system but, at best, to explain the evolution of housing and financial asset prices by means of portfolio decisions affected by psychological factors.

Fifth and finally, following the financial crisis that broke out in 2007–08, a certain amount of interest has emerged among economists and bank regulators in what is called 'network theory', according to which the balance sheets of banks and financial institutions, and also those of non-financial businesses, are interconnected; in other words, most of their assets consist of liabilities issued by other financial institutions. If the value of some assets falls, then insolvency arises and spreads through the system (see European Central Bank, 2010). Insolvency occurs therefore when the value of assets in a balance sheet is less than the liabilities. In this situation, a company cannot repay its liabilities from its assets.

The theory arises from the very obvious observation that, when there are interbank markets (in which banks borrow and lend reserves in order to maintain their ability to make payments to their customers), banks end up having a lot of lending to – and borrowing from – other banks on their balance sheets. A bank that is unable to make the payments on its loans from another bank can then easily transmit its difficulties to other banks. This was certainly a factor in the financial crisis that erupted in 2007–08.

However, precisely because of these interbank markets, banks' balance sheets change very quickly from day to day. So network theory may be useful for simulating how a financial crisis arises. But any model built up using real variables very quickly goes out of date, because the balance sheets of intermediaries change not only with the business, but also owing to financial and monetary innovation, or the emergence of new kinds of financial instruments, and 'shadow' banking, which lies beyond the control of regulators and policy-makers.

The heterodox perspective

What is wrong with the textbook approach?

In general, since the 2007–10 financial crisis, there has been a general agreement regarding the importance of finance in explaining the way in which the economy works, as well as a widespread sense that the explanations commonly found in economics textbooks, usually based upon some particular insight that is obvious, give a very poor account of finance in the economy.

For example, wealthy people make portfolio choices; people save for retirement; there is uncertainty about future stock market prices as well as lack of information about the future of the economy; stock market booms make stockholders wealthier; bank and corporation balance sheets are interconnected. But without an explanation of how the economy as a whole works, these insights remain trivial.

There are at least three reasons why these textbook approaches cannot develop a theory of how the economy as a whole works. First, mainstream approaches reduce all economic activity to the economic choices of individual agents, or households that are making consumption and saving decisions. This therefore limits consideration to arguments and decisions that are made only by households; any attempt to deviate from this theory is held to lack microfoundations. To any economist who believes that all private sector economic and financial decisions are made solely by households, any other theory postulating decision-making by firms or banks would also lack plausibility, when theory is supposed to invite individuals to imagine how they would make economic and financial decisions.

In reality, we cannot conflate all agents together. Households make consumption decisions, with a relatively few wealthy individuals making portfolio investment decisions as well. Firms make production and productive investment decisions upon which we all depend, and a third type of agents, that is, financial intermediaries, make decisions that affect the structure of the financial system.

Second, textbook approaches confuse saving, or savings, with finance and credit. Saving is simply the flow of income that is unspent in a given year or period of time. Savings is the stock of financial assets that exists in an economy at any one time. It consists of the saving that has been accumulated over a period of time. But it also includes a huge amount of credit that has been advanced on the security of assets (financial and non-financial assets, such as housing, consumer durables, factories, land, and so on). Thus, a bank deposit may be backed in its balance sheet by a loan from the bank to another bank. The second bank's loan (a liability) may be backed by an equivalent asset in the form of a loan to, say, a hedge fund. That hedge fund's liability to its bank may have its asset counterpart in the form of shares of an insurance company. The insurance company may back its shares with assets in the form of shares in an industrial company, or government bonds, and so on.

The financial system therefore does not just consist of 'pure' intermediation (transfer of saving) between households with surplus income (saving)

and households with financial deficits ('negative saving', whose expenditure exceeds their income), except in a very trivial sense. The system consists of a more complex intermediation in the first place between households, but also firms and governments. Adding even further complexity to the system is a more advanced kind of intermediation between financial intermediaries to enable them to match the assets and liabilities in their balance sheets by term (the period of time until a loan or financial contract has to be settled) and by liquidity (the ease with which long-term financial assets, or illiquid assets, such as real estate, may be converted into money to make payments on liabilities).

For instance, financial intermediaries, such as hedge funds, private equity or money market brokers, exist to allow wealth holders to have a diversity of assets and liabilities, according to the financing preferences of particular kinds of firms, households and governments.

The most common example of this is a bank that holds short-term deposits, because people prefer to have deposits that can be withdrawn quickly. However, borrowers commonly prefer to borrow for more extended periods of time: years, or even decades. A bank therefore not only collects deposits for lending onto borrowers, but also engages in 'maturity transformation', using short-term deposits to finance long-term loans. Similarly, pension funds and insurance companies hold long-term bonds and shares, which may become illiquid (less easy to sell) and of uncertain value. Private equity funds emerge as an additional kind of financial intermediary that will guarantee the liquidity and profitability of company shares.

This complexity of financial intermediation is almost completely ignored by the textbook approach to financial economics. Yet it has one very important consequence for that key variable in macroeconomics: the rate of interest. Most economists, following Knut Wicksell and Keynes, consider the rate of interest as determined in the money market by the demand and supply of money (nowadays in the form of bank reserves). But, in more recent macroeconomics textbooks, the rate of interest is the factor that brings saving (unspent income) equal to investment by the 'law of supply and demand', with a downward-sloping demand curve for saving to be used for investment purposes, and an upward-sloping supply curve of saving. In other words, most textbooks still adhere to the concept of loanable funds used for investment.

In the real world, however, there is a whole constellation of interest rates, or rates of return on financial assets, as stipulated by Joan Robinson (1952). In that world, these rates of interest bring nothing into equilibrium. Rather,

imbalances between the preferences of depositors and borrowers are met by another layer of intermediation that will satisfy liquidity and term requirements, rather than by adjustment of these requirements by some rate of interest, in accordance with the law of supply and demand. For instance, there is no rate of interest that will make bank depositors with cheque book accounts prefer to hold long-term deposits that will match the long-term loans that borrowers prefer; or that will make borrowers willing to finance their activities with short-term loans. The margin between short-term and long-term rates of interest is supposed to compensate banks for any difficulties arising from their maturity transformation, and banks will borrow and lend among each other to make up any shortfall in money to pay out deposits.

Rates of interest exist to determine the distribution of income or returns around the financial system, or financial surplus among firms, rather than matching needs to provision of finance or money.

In other words, differences in financing requirements and financial investment preferences are accommodated in the financial system not by a rate of interest that brings them into equilibrium, but by new forms of expanding financial intermediation. Differences in the various financial market interest rates exist to make such further forms of intermediation worthwhile.

Third and finally, reality is much more complex than the naïve and simplistic narratives put forward in the textbook approaches to finance. Moreover, underlying financial flows are business fluctuations. As a result, the financial system is unstable and things are constantly going wrong in terms of unconfirmed predictions and unwanted losses in finance. Financial economics attributes all these untoward events in the system to 'risk', this being the standard cliché with which bankers or financiers, when asked to explain events on which they are supposed to be expert but which they do not understand, claim mystical insight into unexpected events that have befallen the markets. The mathematical approach to financial economics attributes a probability distribution to this risk. This has given rise to a whole cottage industry of mathematical modelling of risk in financial markets. Modelling risk does not make losses any more predictable. But it does help to allay anxieties among financial investors (Box 6.3).

The financing needs of modern capitalism

Modern finance emerged out of the financing needs of production and exchange in a capitalist economy. Merchant capitalists in the seventeenth and eighteenth centuries raised 'circulating capital' to finance their cargoes of

Box 6.3

FINANCE: KEYNES VERSUS THE 'CLASSICS'

The economists whom Keynes called the 'Classics' and mainstream economists today argue that the financial system is the system that brings into equilibrium the demand for money for business investments in plant, machinery and buildings, and the supply of loanable funds from savers. The returns from the investments are supposed to pay costs (the income to the lenders) of the loanable funds used to make the investments. However, in Keynes, and even more so in our times, most private sector investment is financed using the financial reserves or savings of the businesses making the investments, and the rest is financed by bank borrowing. Having spent their savings on buying plant, equipment and premises, those businesses will then issue a bond, or new shares, and the money raised in this way is used to replace the depleted savings or the reserves of the investing business, or to repay their bank borrowing. Hence, for Keynes, the business of the financial system is the management of liquidity rather than the financing of investment.

goods, which they sold around the world. With industrial capitalism, financing needs became massive. In addition to materials and labour, finance was needed to buy or build machines and factories. In the seventeenth and eighteenth centuries, special laws had to be passed to allow canal companies, and later railway companies, to issue shares or bonds to finance the construction of these large undertakings. These stocks were traded on stock exchanges alongside government bonds. But in manufacturing, for example, the expansion of production and investment was limited by the amount of money that an individual capitalist possessed (his own capital) and the amount of money that he could borrow from banks, usually on a short-term basis, or obtain from partners.

However, things changed when, in the 1860s and 1870s, the United Kingdom, followed by most countries in Europe and North America, passed Companies Acts. These revolutionized company finance in two ways. First of all, they allowed capitalist businesses to become joint stock companies with limited liability without the trouble and expense of getting the legislature (Parliament or Congress) to pass a special law for each company. All that a business had to do was to have its rules, or Articles of Association, approved by a court of law or Company Registrar. This allowed companies to obtain long-term finance much more easily. Manufacturing companies went from being financed from personal wealth and short-term borrowing from banks, to being financed with long-term bonds or shares, titles of ownership that

gave owners a share of the profits of the company. Of course, financial investors would normally be reluctant to tie up their money for a long time, let alone permanently in the case of shares, with one particular company. A neat arrangement with this new kind of company financing was that any holders of a company bond or share, if they wanted their money back, did not need to demand it back from the company, but could sell their bond or share in the stock exchange to another financial investor. This made it much more attractive to provide long-term finance for companies, because the provider could usually get their money back quickly and easily.

The second consequence of this new type of capitalist financing was that it was now possible for companies to match the term of their capital expenditure to their financing. For example, if a company was financing a factory that was expected to generate profits over a 30-year period, the company could finance it with a 30-year bond, a liability whose interest and capital repayments could be paid out of those profits over the lifetime of the capital asset, the factory, that was being financed. This was a huge advantage to capitalists, who previously would have had to limit their investments to what they could finance themselves, or to money that they could borrow short-term. Such short-term borrowing would have to be rolled over (new borrowing undertaken to repay old borrowing) many times over the course of a capital asset's lifetime. Financing costs, in the form of the interest charged on the borrowing, would be uncertain, and if there was a sudden squeeze on credit and banks became reluctant to lend, capitalists who were unable to repay their debts quickly would be driven out of business. To avoid this kind of crisis, capitalists would have to hold large amounts of money in bank deposits, and this would further restrict investment.

The economic consequences of long-term finance

The easing of financing conditions through 'financial innovation' might have been expected to give rise to a huge investment boom, with rising output and employment. Instead, one by one, capitalist countries, after an initial phase of industrial development, succumbed to economic stagnation and unemployment that, in one way or another, lasted until the middle of the twentieth century.[1] The reason for this was identified initially by the American economist and social critic Thorstein Veblen (1857–1929), and subsequently by the Austrian Marxist Rudolf Hilferding (1877–1941). The ease of issuing long-term bonds and shares, and trading them in the stock market, also facilitated a rapid concentration of ownership of business. The market for long-term capital also created big business or monopoly capital: firms that were so large that they dominated their markets and suppliers.

So what caused this stagnation? For Veblen, an 'underconsumptionist', economic stagnation arose because, with mass unemployment, wages and employment income were too low to buy all the potential output of the economy. Big business played an important part, because large corporations, with access to the capital market, would stimulate a stock market boom, which in turn would arouse a sense of prosperity, a 'wealth effect' for shareholders who would spend on purchasing luxury goods that would result in a temporary boom. However, in the end, luxury consumption was unable to overcome the gap between employment income and the total value of labour: workers could never be paid the full value of their labour, because that would wipe out profits. The deficiency of total demand in the economy would frustrate the realization of expected profits, on which the stock market boom depended. The stock market would crash, and the economy would then revert back to its 'natural' condition of stagnation and mass unemployment (Veblen, 1904).

Rudolf Hilferding had a much clearer idea of how long-term debt markets affected the structure of the capitalist economy. In the first place, monopoly capital was reinforced by coordination with the banking system. He called 'finance capital' this alliance between finance and monopoly capital. Banks would facilitate the export of capital (foreign loans) to assure demand abroad for the output of the monopolists. By coordinating production and forming cartels, the monopolists would be able to stabilize their markets and production. However, this left another group of capitalists – namely smaller, competitive firms without access to long-term debt markets – whose profits depended on competition with the monopolists. Their precarious position in product markets destabilizes capitalism.

Veblen and Hilferding established the idea that financial operations lie at the heart of the operations of the large business corporations that determine the macroeconomics of modern capitalism. This stands in stark contrast to the fundamental postulate of textbook financial economics and finance theory, which argues that the financial system consists of the saving activity of households engaged in the exchange of 'endowments' or domestic production, where firms do not exist and production using large-scale equipment plays no part.

Veblen and Hilferding were then followed by Michał Kalecki (1899–1970) and Josef Steindl (1912–93). Kalecki, a Polish economist, took up Hilferding's analysis to show that the key factor in determining output and employment in the economy is business investment. Investment also plays a critical role in allowing firms to realize profits.

If what capitalists pay their workers (wage income) comes back to capitalists when workers spend their incomes on consumption, then whose expenditure allows capitalists to realize their profits in money form? Kalecki's (1966) answer was that capitalists do this themselves by their expenditure on their own consumption and on business investment.

The idea originates in the schemes of capitalist reproduction that Marx (1885) put forward in Volume II of *Capital*. It came down to Kalecki through the work of Hilferding and Rosa Luxemburg (1870–1919). Kalecki showed how monopoly capital makes the economy even more unstable. Business cycles become more extreme, because the monopoly profits of big firms allow them to afford to maintain unused capacity in a recession, and the unused capacity discourages the investment that could start off a boom. In a boom, those large firms are likely to overinvest in order to reinforce their dominance in their markets. Small and medium-sized firms, without access to long-term debt markets, have to rely on much more short-term finance, getting themselves into debt in order to invest, and then falling victim to excessive debt in a recession.

In Kalecki's theory, a very particular part is played by the saving of wealthy individuals who do not participate in capitalist production, but live off the income (interest and dividends on shares) that they receive from capitalist enterprise. The income of these rentiers is, of course, paid out of the sales proceeds of capitalist firms. It is a portion of the expenditures of firms that does not come back to capitalist producers if those rentiers save their income, instead of spending it on their own consumption. Kalecki pointed out that this rentiers' saving serves eventually to depress production and employment, because firms that do not get back in sales revenue all that they spend on costs of production and financing costs (the incomes of the rentiers) have to borrow the shortfall from the banks in which the rentiers have deposited their saving. In this way, banks' balance sheets (deposits and loans) expand, but productive capitalists get more and more into debt. Eventually this indebtedness makes those capitalists cut back on their investment, and this causes a fall in realized profits, and further reductions in investment, output and employment. In Kalecki's theory, finance is integrated in the analysis of the firm and the household in the unstable macroeconomics of capitalism.

Josef Steindl developed Kalecki's theory further to show that saving is not just characteristic of highly paid rentier capitalists. Steindl argued that the twentieth-century advanced capitalist economy also contains a large middle class consisting of members of the liberal professions (doctors, lawyers, journal-

ists, and so on), government employees, teachers and public service workers, as well as those employed in financial institutions and the managerial bureaucracies of big business. One way or another, their incomes are derived from income generated in production. If they do not spend overall as much as they receive in income, productive capitalists will be forced into financial deficit, unless capitalists can offset this with their investment. (It is assumed here that capitalists cannot sell abroad, and that the government budget is balanced.) The financial deficit has to be financed by borrowing. So household saving in general must be exceeded by business investment if firms overall are to obtain profits (Steindl, 1989). Like Kalecki, Steindl shows how household saving, far from being automatically matched by investment, as in textbook financial economics, can depress firms' profits and cause recession and unemployment.

Finance in Keynes's analysis

Finally, we get to John Maynard Keynes, another economist who was to incorporate long-term debt markets into the foundations of his macroeconomics (see also Chapters 3 and 4 in this volume). The key text in which he did this was his most important book, *The General Theory of Employment, Interest and Money*, published in 1936. The long-term debt or capital markets were responsible for setting the long-term rate of interest, that is, the rate of interest on long-term bonds. This is determined by the prices of bonds, which in turn are determined by the supply and demand for bonds. A bond has a rate of interest fixed when it is issued. If the price of the bond goes up, after the bond is issued, then the yield (or the income that may be obtained from a given amount of money invested in the bond) will be reduced. If the price of the bond falls in the market, then the yield, or expected income from the bond, will go up.

The market yield at any one time sets the rate of interest that industrial capitalists must offer if they are to get long-term financing for their investment. According to Keynes, industrial capitalists, or entrepreneurs, calculate the rate of return that they may expect to receive on their investments, and compare this with the long-term rate of interest. If this expected rate of return on business investment is higher than the long-term rate of investment, then it will be worthwhile investing in these projects. Accordingly, this long-term rate of interest and business expectations of profits determine together the amount of investment in any period in the economy (see Keynes, 1936, Ch. 11; Davidson, 2002, Ch. 7). In turn, as with Kalecki, the amount of investment in a given period determines the amount of output and employment in an economy.

According to Keynes, there are two flaws in the way this arrangement works in a capitalist economy. First of all, the financial investors in the stock market are not truly committed to the enterprises in which they invest. They can make more profit from speculation – that is, buying bonds or shares in anticipation of an increase in their price – than from the dividends paid on these shares, or interest on the bonds. Such speculation is done by anticipating opinion in the stock market, rather than the profits of companies. This makes the stock market an unreliable source of finance: 'When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done' (Keynes, 1936, p. 159). The second flaw is that expectations of future profits, after payment of interest, are based on subjective rather than objective knowledge. Such expectations are therefore volatile, and the resulting shifts in investment are the cause of business cycles (Keynes, 1936, Ch. 22).

Keynes's solution for this was to use monetary policy to drive down interest rates, including the long-term rate of interest, so that business investment was maintained at a level that would ensure full employment. This policy he called the 'euthanasia of the *rentier*' (Keynes, 1936, p. 376). Fiscal policy – that is, government expenditure – could also be used to stabilize aggregate demand.

In the end, finance will always escape the simple definition of its relationship with the rest of the capitalist economy. This is because the economy itself is unstable, and finance is merely a part of that instability. Devoted to demonstrating equilibrium, mainstream finance and macroeconomics do not have the necessary concepts or analysis with which to understand – let alone explain – that instability. But the theorists who did were those who thought in terms of business cycles. This is why Keynes, Kalecki, Schumpeter, Hilferding, Steindl and Minsky (an economist whose portrait we present at the end of this chapter) are better guides to financial macroeconomics than those who write textbooks on financial macroeconomics.

Modern finance

Modern finance has moved some way since the days of Keynes and Kalecki, let alone Veblen and Hilferding. In the twenty-first century most stocks and shares are owned by pension funds and insurance companies, rather than individual rentiers, and most of the value of bonds and stocks in existence is represented by bonds issued by governments to finance their expenditure. Pension funds and insurance companies usually have their portfolios, or holdings of financial assets, regulated in accordance with their liabilities (their future

pension or insurance pay-outs). They tend to hold stocks for long periods, because turning over their portfolios in the stock market has high transaction costs, for example in brokerage fees. Most of the activity in stock markets is therefore done by more marginal holders of stocks and bonds, hedge funds and private equity firms, sometimes referred to as 'shadow banks' because they are not strictly regulated like banks, pension funds and insurance companies. These shadow banks follow complex financial investment strategies to make speculative gains from the markets in financial assets.

The line between pension and insurance institutions, on the one hand, and shadow banks on the other hand, is not clear-cut. When interest rates on normal bonds are low, these institutions have an incentive to seek out speculative gains by delegating management of portions of their portfolios to private equity firms or hedge funds. This incentive has been increased by the near zero interest rates since the financial crisis of 2007–08, and the buying up of bonds by central banks under quantitative easing policies.

Conclusion

Governments and companies benefit from having a financial sector that allows them to finance themselves by issuing long-term bonds and shares. These have the advantage of fixing the money cost of finance and saving on cash resources when borrowers with long-term financial commitments are not obliged to keep rolling over short-term bank debt. However, the secondary markets where already issued stocks and shares are traded are disturbed by speculative activity, and the underworld in a system dominated by large, capital-market financed corporations is one where small and medium-sized firms, without access to capital markets, are forced into debt because corporations prefer speculative gains to productive investment.

Most economists have difficulties in understanding the network of debts that arise with the emergence of long-term finance. Most commonly these economists confuse such debt structures with savings, and ardent reformers among them would like finance to be confined to providing resources for productive investment. However, the financial system is really about liquidity that allows governments and firms to have access to long-term finance, at the same time as allowing the institutions that provide that finance to value their portfolios and withdraw from particular financing commitments without affecting the firms and governments who have obtained that finance. The way to make finance safe for firms, governments and citizens is for governments and central banks to regulate that liquidity through debt management and the buying and selling of government securities.

NOTE

1 This does not mean that there were no new industries, most notably those based on the internal combustion engine (cars, ships, and so on), chemicals and electronics. But these were usually balanced by declining incomes and employment in agriculture and railway construction. See Heilbroner (1961).

REFERENCES

Bernanke, B.S. and M. Gertler (1989), 'Agency costs, net worth, and business fluctuations', *American Economic Review*, **79** (1), 14–31.

Davidson, P. (2002), *Financial Markets, Money and the Real World*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Diamond, D.W. and P.H. Dybvig (1983), 'Bank runs, deposit insurance and liquidity', *Journal of Political Economy*, **91** (3), 401–19.

European Central Bank (2010), *Recent Advances in Modelling Systemic Risk Using Network Analysis*, Frankfurt am Main: European Central Bank.

Heilbroner, R. (1961), *The Worldly Philosophers: The Lives, Times and Ideas of the Great Economic Thinkers*, New York: Simon & Schuster.

Kalecki, M. (1966), 'Money and real wages', in *Studies in the Theory of Business Cycles 1933–1939*, Oxford: Basil Blackwell, pp. 40–71.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Marx, K. (1885), *Das Kapital, Volume II*, Hamburg: O. Meissner.

Modigliani, F. and M. Miller (1958), 'The cost of capital, corporation finance and the theory of investment', *American Economic Review*, **48** (3), 261–97.

Robinson, J. (1952), *The Rate of Interest and Other Essays*, London: Macmillan.

Shiller, R. (2000), *Irrational Exuberance*, Princeton, NJ: Princeton University Press.

Shiller, R. (2003), 'From efficient markets theory to behavioural finance', *Journal of Economic Perspectives*, **17** (1), 83–104.

Steindl, J. (1989), 'Saving and debt', in A. Barrère (ed.), *Money, Credit and Prices in Keynesian Perspective*, London: Macmillan, pp. 71–89.

Stiglitz, J.E. and A. Weiss (1981), 'Credit rationing in markets with imperfect information', *American Economic Review*, **71** (3), 393–410.

Veblen, T. (1904), *The Theory of Business Enterprise*, New York: Charles Scribner's Sons.

A PORTRAIT OF HYMAN PHILIP MINSKY (1919–96)

Hyman Philip Minsky came from a poor family in Chicago in the United States to study mathematics at the University of Chicago in the years preceding the Second World War. There, he met Polish Marxist Oskar Lange, who persuaded him that if he wanted an alternative to the Great Depression affecting the United States, he should study economics.

Minsky graduated in economics and, after military service and a short period working on Wall Street in New York, went on to do a PhD under Joseph Schumpeter at Harvard University. After Schumpeter's death in 1950, his supervision was taken over by Wassily Leontief. Minsky's PhD thesis argued that business cycles were also affected by debt in the economy.

Over the following four decades, Minsky developed a theory of financing structures in the economy, where an investment boom would be accompanied by growing indebtedness, which would cause the boom to break down into recession when income was insufficient to service debts. Debt structures would be 'hedged' if payments on them were matched by income flows. However, all business investment went through a period when income fell below debt service payments. He called such financing 'speculative'. Finally, if debt service payments could only be met by additional borrowing, this was 'Ponzi' finance, named after a well-known Boston pyramid banking operator at the end of the First World War. Minsky argued that in any economic boom, financing structures would deteriorate, with hedged financing becoming speculative, and speculative financing eventually becoming Ponzi. This would continue until a financial crisis broke out. He called this theory 'the financial instability hypothesis'.

In this way, Minsky solved the problem of where financial risk comes from. At any one time, in an indebted capitalist economy, that economy would require a certain level of investment to generate the income necessary to service these debt structures. The problem was that eventually the build-up of debt would induce firms to reduce their investment. This would then cause income, or firms' sales revenues, to fall short of what was necessary to service debts among firms in the economy. Minsky famously suggested that 'stability is destabilizing', because debt commitments increase during periods of economic and financial stability. He also saw this in terms of the liquid assets (bank deposits and short-term financial assets) that companies hold in case of unexpected bills. Minsky thought that, in an extended period of economic stability, companies, banks and financial institutions would come to expect good cash flows as a permanent or even natural state of the economy. They would then reduce their holdings of liquid assets, so that even a small reduction of income, or an unexpected expense, could leave them without the means to pay their bills, bringing on a financial crisis.

Minsky therefore correctly recognized that businesses collapse not because they have failed to maximize profits, as economists think, but because of illiquidity: because they have their financing tied up in assets that cannot be sold quickly to make urgent payments.

EXAM QUESTIONS

True or false questions

1. The financial system allocates saving to investment.
2. The central bank's rate of interest decides how much investment is carried out in an economy.
3. Private equity firms speculate.
4. Borrowing from a bank is more advantageous compared to issuing bonds.
5. A portfolio contains an investor's assets.
6. Ponzi finance is a way of stabilizing the financial system.
7. The stock market allows any company to raise long-term finance.
8. The stock market allows any company to raise short-term finance.
9. Middle-class saving is a way of stabilizing the financial system.
10. Corporations are financed with short-term borrowing.

Multiple choice questions

1. Keynes thought that the stock market is:
   a) a good thing;
   b) a bad thing;
   c) a casino;
   d) a brothel.
2. Hedging provides:
   a) a home for wildlife;
   b) a stable debt structure;
   c) a nice balance sheet;
   d) work for accountants.
3. The size of the US financial system is:
   a) just right;
   b) too large;
   c) too small;
   d) due to borrowing on financial securities.
4. 'Asymmetric information' means that:
   a) the lender does not know what the borrower will do with the loan;
   b) the borrower does not know what they will do with the loan;
   c) banks know who is a good credit risk;
   d) banks know that you do not need the money.
5. Most financial securities are owned by:
   a) the elite;
   b) the Chinese government;
   c) pension funds and insurance companies;
   d) central banks.
6. Shadow banks operate:
   a) in the shadows;
   b) in Monaco;
   c) in foreign exchange markets;
   d) in capital markets.
7. The long-term rate of interest is:
   a) the market yield on bonds;

b) the dividend paid to shareholders;
c) the risk-free return from speculation;
d) the interest offered by banks.

8. A rentier is:
a) someone who derives their income from financial investments;
b) a pensioner;
c) a speculator;
d) someone who borrows from the government.

9. The stock market exists to:
a) allow firms to borrow from banks;
b) allow banks to borrow from central banks;
c) allow governments to buy shares;
d) allow pension funds to sell shares.

10. Financial crisis is caused by:
a) shocks;
b) stability;
c) too much debt;
d) too little debt.

7

Central banking and monetary policy

Louis-Philippe Rochon and Sergio Rossi

OVERVIEW

This chapter:

- explains the essential role of a central bank, which is the settlement institution for banks that, together with the central bank, form a banking system, that is, a system characterized by money homogeneity and payment finality for all its members;
- shows that the central bank, like any bank, is both a money and a credit provider, as it issues a means of final payment and provides credit to the banking sector, which needs both to avoid financial crises;
- presents monetary policy strategies, instruments and transmission mechanisms in an endogenous money framework, characterized by the fact that money supply is credit-driven and demand-determined;
- argues that monetary policy goals should go beyond price stability, to include also financial stability and macroeconomic stabilization, with regard to both economic sustainability and employment levels.

These points are relevant to understanding the scope and limits of monetary policy, and to critically discussing both contemporary monetary policies and the regulatory reforms carried out at national and international levels in the aftermath of the global financial crisis that erupted in 2007–08.

KEYWORDS

- **Bubble:** A steady increase in asset prices on financial and/or real estate markets that cannot be explained by the economic performance of a country, but results from some forms of speculation induced by banks' behaviour in providing credit.

- **Final payment:** A payment as a result of which the payer settles their debt against the payee, who has thereby no further claims on the payer. This provides for monetary order and is a necessary, though not sufficient, condition for financial stability.
- **Monetary targeting:** A monetary policy strategy targeting a publicly announced rate of growth for a given monetary aggregate, in order to make sure that the price level is stable over the long run.
- **Money multiplier:** The ratio between the total money supply and central bank money, which orthodox economists believe exists due to their view of the supposed exogenous supply of the monetary base.
- **Quantitative easing:** A monetary policy intervention whereby the central bank buys a very large volume of financial assets (corporate bonds, government bonds, and so on), in order to lower rates of interest and thereby support economic growth.
- **Taylor rule:** A stylized rule that central banks might use when setting their policy rates of interest, considering inflation and output gaps according to their reaction function against the current or expected macroeconomic situation.

Why are these topics important?

This chapter is important to understanding the role of the central bank in a monetary economy of production as well as the scope and limits of monetary policy-making in such a framework. The large majority of the academic literature and of monetary policy interventions fail to consider the particular nature of money and the actual working of a monetary production economy, with the result that not only are these interventions unable to deliver the results that are expected, but they may also in fact make the economic situation worse.

This chapter will offer a deep criticism of mainstream monetary policy, and in doing so will discuss the so-called impotence of monetary policy, a topic that has been revived of late by Lawrence Summers, the former United States Treasury Secretary and Chief Economist of the World Bank, though the notion goes as far back as Keynes, who once claimed, '[i]t is to the theory of a generalised monetary economy, i.e. of an economy in which, through the fault or the inaction or the impotence of the monetary authority . . . that this book will attempt to make a contribution' (Keynes, 1979, pp. 67–8). This chapter is written in this spirit.

Further, this chapter provides a systemic view of money and banking, enabling us to understand the monetary–structural origins of the global financial crisis that erupted in 2007–08, particularly after the demise of Lehman Brothers in the United States. It thereby offers an explanation of this crisis based on

structural rather than behavioural factors, which the regulatory reforms put into practice at national as well as international levels do not and cannot address. This chapter thus highlights the largely ignored fact that a monetary–structural reform of domestic payments systems is required in order to eradicate the factors of systemic crises, which are the hallmark of a monetary disorder affecting the working of our national economies. As a result, the objectives of monetary policy should go much beyond price stability, contributing to re-establishing monetary order and guaranteeing financial stability at systemic level.

To summarize, this chapter will consider: (1) monetary policy and its transmission mechanism from a mainstream and a post-Keynesian perspective; and (2) the architecture of monetary policy.

The mainstream perspective

The mainstream approach to central banking and monetary policy focuses on the relationship between monetary policy and inflation. In its original form, this meant that there existed a direct relationship between the growth of the money supply and the rate of inflation. This rested on the notion that central banks could effectively control changes in the money supply, which American economist Milton Friedman likened to 'helicopter money' (Friedman, 1969, pp. 4–5), and that in turn these changes would impact upon the price level. This statement carries many implications. For instance, it means that the central bank is in a position to steer and determine the total amount of money within the economic system, either directly or indirectly. Directly as regards the so-called 'high-powered money' (see Friedman, 1969), which is the monetary base issued by the central bank (Box 7.1). Or indirectly, via the so-called 'money multiplier', that is, the relationship that orthodox economists believe exists between the monetary base and the total amount of money within the system, through bank lending (Box 7.2).

This mainstream view gave rise to specific policies of monetary-targeting strategies by a number of central banks following the collapse of the Bretton

BOX 7.1

THE MONETARY BASE

The monetary base includes cash in circulation as well as the deposits of commercial banks at the central bank, also known as reserves. It is assumed that central banks can effectively control the monetary base, increasing or decreasing it at will.

BOX 7.2

THE MONEY MULTIPLIER

This theory is closely linked to the central bank control over the monetary base, and sees a relationship between the monetary base and overall money supply in the economy, through the lending activities of commercial banks. In essence, the money supply is a 'multiple' of the monetary base. When central banks increase the reserves of commercial banks (an exogenous action), the latter could then use these reserves and increase lending to the private sector, which then creates money. Moreover, this multiplier is assumed constant, so that whenever the central bank increases the monetary base, it translates into a predictable and proportionate increase in the money supply. Since the central bank is said to control the monetary base, it is further assumed that it can also control the growth of the money supply.

BOX 7.3

THE BRETTON WOODS REGIME

Established in 1944, the Bretton Woods Agreement, so called because the meetings were held in Bretton Woods, a small town in New Hampshire in the United States, was meant to establish the rules governing monetary relations between countries, and in particular their exchange rates, that is, the relative price of the currencies between countries, which had to keep the value of their respective currencies stable relative to the price of gold.

Woods regime in the early 1970s (Box 7.3). Indeed, in the 1980s and 1990s, the majority of central banks implemented their monetary policy with a view to reaching some previously announced target for the rate of growth of a given monetary aggregate; as M0, that is, central bank money, or as M3, that is, the total sum of bank deposits across the banking system, including both the central bank and the set of commercial banks (Box 7.4).

This assumed control over the money supply is at the heart of the original mainstream monetary model, which was based on the so-called 'quantity theory of money', which further assumed a direct influence of the money supply on inflation (Box 7.5).

In other words, according to this view, central banks could effectively influence inflation directly via a proportional relationship that this theory establishes

BOX 7.4

MEASUREMENTS OF MONEY SUPPLY

Economists have designed several aggregates through which they aim at measuring the money supply. Here is a short list, along with what they measure:

- M0: total bank reserves; it is also referred to as 'narrow money' or the monetary base.
- M1: currency, travellers' cheques, bank accounts such as sight and demand deposits, and other checkable deposits; it excludes reserves, bonds and savings accounts.
- M2: this is a broader definition than M1, and also includes savings deposits, money market securities, mutual funds and other types of time deposits.
- M3: this is an even broader definition than M2, and includes less liquid assets, and as such sees money more as a store of value.

Central banks and economists have at various times used any of these definitions of monetary aggregates to see whether they contained any information regarding possible measures of inflation. M3 in particular was used for many years for such a purpose. Today, central banks have largely abandoned M3 in favour of another definition of money, called MZM (money zero maturity), which measures money's liquidity in the economy. It includes cash (bills and bank notes), chequing and saving accounts, as well as money market accounts.

BOX 7.5

THE QUANTITY THEORY OF MONEY

According to the quantity theory of money, the price level in the economy is directly related to the amount of money in circulation, or the money supply. Hence, any increases in the money supply would lead to an increase in prices (inflation), because there would now be a situation of 'too much money chasing too few goods'. Dating as far back as the sixteenth century, this theory became an important component of conservative British philosopher David Hume. The theory was later revived by Milton Friedman.

between (changes in) the money supply (M) and (changes in) the general price level (P), according to the following so-called 'equation of exchange':

$$MV \equiv PQ \tag{7.1}$$

where V represents the so-called 'velocity of circulation' of money (the number of times that a unit of money is supposed to be used in payments across the

economic system during a given period, say a calendar year) and Q is physical output (the bulk of goods and services produced during the same period of time). This equation of exchange states that the value of money times its velocity (left-hand side) is equal to the value of output (right-hand side). This suggests therefore that the money supply and the price level are connected.

Though developed very early in the nineteenth century, the equation of exchange was revived in the 1960s by Milton Friedman of the University of Chicago, who insisted on reading into it a causality from the left-hand side to the right-hand side. In doing so, Friedman suggested that money caused prices or, in terms of changes, a change in the money supply caused a change in the price level, that is, inflation. This led Friedman (1987, p. 17) to pronounce his famous saying: '*inflation is always and everywhere a monetary phenomenon*'. This was the mantra of the monetarist school (Box 7.6).

BOX 7.6

MONETARISM

Monetarism is a school of thought developed by Milton Friedman, as a direct challenge to Keynesian economics, which rested on the use of fiscal policies. Friedman developed these views in particular in a monumental book, co-written with Anna Schwartz, entitled *Monetary History of the United States 1867–1960*. In this book, Friedman and Schwartz (1963) argued that, historically, excessive growth of the money supply contributed to periods of high inflation. An adversary of Keynesian economic policies, Friedman rejected the importance of fiscal policy and instead favoured the use of monetary policy. Monetarism emphasizes that changes in the money supply have a direct and predictable impact on inflation. This approach is based on the quantity theory of money, which stipulates that $MV \equiv PQ$. Rewriting this equation in terms of growth rates, we have $\dot{M} \equiv \pi + g - \dot{V}$. Monetarists assumed that the growth of the velocity of money was zero, so that changes in the money supply caused changes in the price of output. Over the long run, it is possible to find a so-called natural rate of economic growth. This being the case, if central banks wanted an inflation rate of, say, zero, the proper policy would be to allow the money supply to grow at the same rate as long-run economic growth (fixed monetary rule), thereby eliminating 'excess' money growth. This is why the assumption of exogenous money or control over the growth of the money supply by the central bank is paramount: it was imperative for Friedman to assume it in order to undermine Keynesian economics. As the authors of this book explain, however, money is endogenous, which undermines the essence of monetarism. This said, one does not need central bank control over the monetary base to justify inflation targeting policies inspired by monetarism.

Rearranging the terms of equation (7.1) and expressing it in growth rates form, this determines the factors affecting the rate of growth of the money supply, as in equation (7.2):

$$\dot{M} \equiv \pi + g - \dot{V} \tag{7.2}$$

where a dot over a variable indicates its rate of change over time, π is the measured rate of inflation on the goods market, and g represents the growth rate of produced output. Box 7.6 gives a summary of this view.

Equation (7.2) has been used to set up and justify monetary-targeting strategies, once the central bank was in a position to determine the three variables on the right-hand side of that equation (see Box 7.6). The German Bundesbank, for instance, was implementing this strategy in the 25 years that preceded its adoption of the euro on 1 January 1999 (Table 7.1).

The Bundesbank decided the targeted growth rate of the relevant monetary aggregate (ΔM^*), applying equation (7.3):

$$\Delta M^* \equiv \pi^{normative} + g^{potential} - \Delta V \tag{7.3}$$

where $\pi^{normative}$ is the desired rate of inflation, $g^{potential}$ is the rate of growth of potential output, and ΔV is the rate of change in money's 'velocity of circulation'.

Provided the central bank controls the monetary base and we assume a constant relationship between that base and the money supply, monetarism thus believed in the ability of central banks to target a specific rate of inflation.

While monetarism has today been abandoned by central banks – that is, the notion that central banks control the money supply – the spirit of this approach remains very much entrenched in policy circles. While many or most central banks have now recognized their inability to control the money supply (Box 7.7), they have now favoured an approach that focuses on their ability to control the rate of interest, something post-Keynesians have been saying for several decades now.

Yet, the mainstream version of this interest rate procedure is much different. First developed by American economist John Taylor in a 1993 paper, it stipulates that central banks should now set the interest rate at a level that generates an inflation rate consistent with their targeted inflation rate, or

Table 7.1 The monetary-targeting strategy of the German Bundesbank, 1975–98

Year	$\pi^{normative}$ (%)	$+g^{potential}$ (%)	$-\Delta V$	$=\Delta M^*$ (%)
1975	5.0–6.0	–	–	8.0
1976	4.0–5.0	2.0	Increasing	8.0
1977	≤ 4.0	3.0	Increasing	8.0
1978	3.0	3.0	–	8.0
1979	Moderate	≈ 1978	Declining	6.0–9.0
1980	3.5–4.0	3.0	Declining	5.0–8.0
1981	3.5	2.5	Increasing	4.0–7.0
1982	3.5	1.5–2.0	–	4.0–7.0
1983	3.0	1.5–2.0	–	4.0–7.0
1984	2.0	2.0	–	4.0–6.0
1985	2.0	2.0	–	3.0–5.0
1986	2.0	2.5	–	3.5–5.5
1987	2.0	2.5	–	3.0–6.0
1988	2.0	2.0	–0.5	3.0–6.0
1989	2.0	2.0–2.5	–0.5	5.0
1990	2.0	2.5	–0.5	4.0–6.0
1991	2.0	2.5	–0.5	3.0–5.0
1992	2.0	2.75	–0.5	3.5–5.5
1993	2.0	3.0	–1.0	4.5–6.5
1994	2.0	2.5	–1.0	4.0–6.0
1995	2.0	2.75	–1.0	4.0–6.0
1996	2.0	2.5	–1.0	4.0–7.0
1997	1.5–2.0	2.25	–1.0	3.5–6.5
1998	1.5–2.0	2.0	–1.0	3.0–6.0

Source: Bofinger (2001, p. 251).

BOX 7.7

THE COLLAPSE OF MONETARISM

In the late 1970s and in the 1980s, monetarism was practiced by central banks around the world. But following the oil shocks of 1973 and 1979, inflation rates increased to unprecedented levels. For instance, inflation in the United States climbed as high as 14.76 per cent in April 1980, and interest rates (federal funds rate) to 19.83 per cent on 29 June 1981, and the prime interest rate rose to 21.5 per cent in June 1982. The monetarist experiment proved to be a disaster, and was abandoned, though its fundamental message remains.

some targeted monetary aggregate that in turn would deliver a desired rate of inflation. In this sense, the money supply rule advocated by Milton Friedman was replaced with an interest rate rule. Yet, the core idea of monetarism, that inflation is influenced by central bank policies, remains in practice. With the Taylor rule, inflation becomes 'always and everywhere' a monetary policy phenomenon.

Monetary-targeting strategies have usually been operationalized via an implicit interest rate rule, which can be written as follows:

$$i_t = i_{t-1} + \alpha\,(\mu_t - \mu^*) \tag{7.4}$$

where i represents the policy rate of interest set and controlled by the central bank during the relevant period represented by t, α is a positive parameter, μ_t is the observed growth rate of the targeted monetary aggregate, and μ^* is the targeted rate of growth for that monetary aggregate. We can also assume that central banks target a specific rate of inflation; a strategy that has been called inflation targeting. In this sense, the above equation simply becomes:

$$i_t = i_{t-1} + \alpha\,(\pi_t - \pi^*) \tag{7.5}$$

In equation (7.5), the central bank is now targeting a specific rate of inflation, say 2 per cent. Accordingly, if the rate of inflation is above its target level, or similarly if the monetary aggregate is above its targeted level (as in equation 7.4), the rule dictates that the central bank should increase the rate of interest.

Note that, in both instances, this shows that monetary-targeting central banks have been well aware, in general, of the fact that they cannot steer the relevant monetary aggregate through the 'quantity theory of money' channel. In its stead, they operated through interest rate channels, that is, a mechanism centred on the indirect relation between interest rates and demand on the market for produced goods and services, which in the end affects the price level assuming a full-employment situation.

The question then becomes obvious: how do changes in interest rates filter through and deliver on a targeted inflation rate? In other words, what is the transmission mechanism of monetary policy? The causal chain in this approach is actually a two-step process, in what has been called New Consensus Macroeconomics, that runs like this:

$$\uparrow i \Rightarrow \downarrow C, \downarrow I \Rightarrow \downarrow TD$$

BOX 7.8

THE PHILLIPS CURVE

In 1958, A.W. Phillips published a paper in which he observed an empirical relationship between the rate of unemployment and the increase in nominal wages in the United Kingdom, between 1851 and 1957. In their interpretation of this article, Keynesians argued that there existed a short-run trade-off between fighting either unemployment or inflation, such that fiscal or monetary policies could not deliver low unemployment and low inflation simultaneously.

In the first instance, central banks will change their interest rate (i) as described above, whenever observed or actual values of inflation are above the targeted rate. Specifically, they will increase the rate of interest. Then, in a first instance, this should translate into lower consumption (C) and investment (I).

Second, this drop in total demand (TD) should translate into an increase in unemployment rates (U), which in turn lowers the inflation rate (P):

$$\downarrow TD \Rightarrow \uparrow \mathrm{U} \Rightarrow \downarrow P$$

This second step is crucial for monetary policy, and is known as the Phillips curve (Box 7.8): a theory that describes the short-run inverse relationship between the unemployment rate and the rate of inflation. As unemployment increases, inflation comes down, and vice versa.

Today, monetary policy depends on this two-step process to achieve price stability: increases in the rate of interest should therefore bring inflation rates back down to their targeted level. As suggested above, while this newer version abandons the idea that central banks control the money supply, it retains the idea that inflation is related to monetary policy, and that central banks have an important role to play in achieving price stability.

The above discussion can also be extended in the context of an open economy, that is, an economy that considers both imports and exports, but also the exchange rate. If one considers such an economy, one should expand on this interest rate view by including the effects on total demand and the price level that stem from an exchange rate depreciation (appreciation) induced by a reduction (increase) in the policy rates of interest administered by the central bank. Indeed, it is assumed that if central banks increase the domestic rate

of interest above the world rate, this increases the demand for the domestic currency as investors move their funds into the domestic economy to take advantage of higher returns, thereby pushing up the demand for the domestic currency and hence its value. The inverse applies when domestic rates of interest are below the world rate.

This exchange rate pass-through may be particularly important for small open economies whose output is often more focused on a limited variety of products, rather than large open economies, whose domestic magnitudes, like the rate of inflation, are less sensitive to the exchange rate channel.

Further, the interest rate channel as described above also affects an economic system's financial stability, as it could generate an asset bubble on either financial or real estate markets (in the absence of the proper macroprudential regulations), when the policy rates of interest are kept too low for too long. This induces debtors as well as creditors to increase the volume of credit in order for both to profit, in one way or another, from the inflating credit bubble (see Chapter 6 for analytical elaboration). In this perspective, the mainstream explanation of the subprime bubble and ensuing crisis observed across the US housing market during the late 1990s and early 2000s is based basically on a too-generous monetary policy by the United States (US) Federal Reserve in the aftermath of the so-called 'dot.com' bubble. The bursting of this bubble led the US monetary authority to dramatically reduce policy rates of interest in an attempt to kick-start the domestic economy, which was also negatively affected by the terrorist attacks of 11 September 2001.

In this light, mainstream economists believe that the policy rates of interest decided by central banks, when used in a countercyclical manner, are a powerful instrument with which to fine-tune economic activity and affect total demand and, hence, have an impact on the price level in order to guarantee price stability across the economy. The spending behaviour of both consumers and businesses would thus be determined by a variable – the rate of interest – that central banks could influence, if not control, through a reduction or an increase in their own policy rates of interest. This is particularly so with respect to their credit lines: if a central bank (say) reduces its policy rates of interest, banks will reduce (although not one-to-one) their own rates of interest, thereby inducing more consumers and businesses to borrow from the banking sector in order to expand production when the economic system is affected by a recession (such as in the aftermath of a financial crisis).

However, as explained by Keynes, reducing the policy rate of interest (to close to zero) may not be enough in order to spur economic growth: it is

akin to pushing on a string. As regards firms in particular, their propensity to invest could be impaired by the economic prospects and the high level of uncertainty regarding the expected profitability of their investment, which, in the case of a deep recession elicited by a global financial crisis, may be so dreadful that firms are unwilling to borrow from banks to finance investment even though the borrowing rate of interest is very low. These firms consider that they would not be able to sell the newly produced output because of the bad economic situation as well as its likely evolution over the relevant time horizon (see Chapters 4 and 9).

Once low or near-zero interest rates failed to stimulate economic activity, the mainstream perspective induced a number of central banks, and in particular the US Federal Reserve System in the aftermath of the global financial crisis that erupted in 2007–08, to try to support economic growth with so-called unconventional policies, in particular quantitative easing (QE); that is, an expansionary monetary policy aimed at making sure that banks lend to creditworthy non-financial businesses (private sector firms), in order for these businesses to carry out their investment projects and thus expand economic activity, thereby increasing the level of employment.

Quantitative easing may take different forms, including the purchase by the central bank of government bonds on the primary market (where these bonds are issued to finance government spending) and the easing of credit standards that banks must fulfil in order to borrow from the central bank (with or without a collateral, that is, a series of eligible assets used as a repayment guarantee for a central bank's loans). Be that as it may, quantitative easing increases the volume of central bank money (reserves) that banks could spend on the interbank market. The aim of this policy, according to the mainstream, is that the increase in reserves would translate into higher bank lending and credit to non-financial businesses.

There are several problems with this reasoning. First, banks do not need reserves to lend (see Chapter 5), and as such, more reserves do not lead to more lending. Banks do not lend reserves to the private sector, as it is not their function. So, in that sense, increasing reserves serves no purpose with respect to bank lending. That is the meaning of endogenous money. Banks were never constrained in their ability to lend. This was essentially the conclusion of a Bank of England report, which stated that: 'As a by-product of QE, new central bank reserves are created. But these are not an important part of the transmission mechanism . . . these reserves cannot be multiplied into more loans and deposits and . . . these reserves do not represent "free money" for banks' (McLeay et al., 2014, p. 1).

Second, as a matter of fact, lenders and/or borrowers may have a 'liquidity preference', which means that they decide to abstain from lending or borrowing in light of their own expectations about the future. Indeed, if the economic situation is bad and its expected evolution is similar, banks and their potential borrowers are likely to postpone any expansion of credit activities, thus making quantitative easing ineffective in supporting economic growth. In this sense, banks were never constrained in lending, although they may have been unwilling to do so.

Third, rather than an inflationary pressure, quantitative easing could contribute to deflation, as agents think or observe that such an expansionary monetary policy has no positive effect on prices, which can fall as a result of agents' abstention from both consumption and investment across the economic system.

Further, reducing the policy rates of interest to zero (or even to the negative domain) may not be enough to reduce banks' own rates of interest, particularly on the interbank market. Shortly after the collapse of Lehman Brothers in the United States on 15 September 2008, the US monetary authority reduced the federal funds rate of interest close to zero, but in fact the interbank market rate of interest increased, since banks were reluctant to lend on this market because of their higher uncertainty (notice again the importance of uncertainty) about the actual financial situation of their counterparts on that market. Indeed, the interest rate channel may not work as monetary authorities expect, although a long period of very low policy rates of interest can have various problematic effects in the financial sector and beyond it, in the absence of a properly designed regulatory framework.

For instance, banks could be induced to borrow from the central bank in order for them to inflate a financial bubble, through a sustained increase in the purchase of financial assets, particularly those assets whose high risks may lead many financial institutions to buy them with the expectation of earning high yields. Also, banks and non-bank financial institutions could inflate a real estate bubble, particularly as a number of middle-class individuals may be attracted by favourable borrowing terms in order to become homeowners. As shown by the subprime bubble that inflated during the 1995–2005 period and burst in 2006, spurring a series of systemic effects in the United States and beyond it, a number of non-performing loans can increase banks' financial fragility to a point where several financial institutions become bankrupt. In particular, those financial institutions that are 'too big to fail' – individually or as a group – will require an intervention by the public sector, which must bail them out to avoid a systemic financial crisis that will induce a deep recession, if not a depression (like the Great Depression of the 1930s).

This state of affairs has induced many regulatory reforms at national and international levels, such as the so-called Basel Agreements signed within the framework of the Basel Committee on Banking Supervision, at the Bank for International Settlements. The large majority of these reforms aim to affect the banks' behaviour, to limit their risk-taking attitude and provide them with a cushion of safety that should be solid enough to avert a systemic crisis in the case of financial or real estate turmoil.

In fact, however, these regulatory reforms do not and cannot avoid a systemic crisis, as this is the result of a structural disorder affecting the payments system, rather than stemming merely from agents' behaviour. The mainstream perspective is not in a position to detect and explain such a systemic crisis, because it considers that any economic system is merely the result of a series of demand and supply forces that the set of economic agents exert on any kinds of market. In this view, therefore, it all boils down to individual behaviour, so that the working of the system as a whole is studied (and can be studied appropriately) with a system of equations that allow to establish the system's 'equilibrium', provided that the public sector does not intervene in the 'market mechanism', because this intervention would be a matter of hindrance and not a macroeconomic stabilizing factor.

The heterodox perspective

The above discussion of the mainstream views on monetary policy is open to severe criticism. Before turning our attention to the post-Keynesian or heterodox alternative, let us consider some of these criticisms.

Assuming inflation is above its target, current monetary policy is based on a two-step process: (1) increases in the rate of interest should lead to a decrease in investment and consumption; which (2) should lead to an increase in unemployment and a decrease in inflation (the Phillips curve), back to target.

Yet, there is empirical evidence that shows this process is severally limited. First, regarding the relationship between changes in the rate of interest and their impact on consumption and investment, there is evidence that shows the latter two magnitudes are not very interest-sensitive, as explained by the following quotation:

> The transmission mechanism from monetary policy to aggregate spending in new consensus models relies on the interest sensitivity of consumption. It is difficult, however, to find empirical evidence that households do indeed raise or lower consumption by a significant amount when interest rates change. Some

> authors have generalized the link to include business investments (see Fazzari, Ferri, and Greenberg, 2010 and the references provided therein) but a robust interest elasticity of investment has also been difficult to demonstrate empirically. (Cynamon et al., 2013, p. 13)

Even the US Federal Reserve is sceptical, as argued in the following quotation: 'A large body of empirical research offers mixed evidence, at best, for substantial interest-rate effects on investment. [Our research works] find that most firms claim their investment plans to be quite insensitive to decreases in interest rates, and only somewhat more responsive to interest rate increases' (Sharpe and Suarez, 2014, p. 1). In this connection, Nobel laureate Paul Krugman stated that '[a]ny direct effect on business investment is so small that it's hard even to see it in the data' (Krugman, 2018).

This would suggest therefore that attempts by central banks to manipulate aggregate demand may not work or, at the very least, are ineffective. If these statements are not devastating enough, it is the second relationship, the Phillips curve, that bears even more criticism.

Claudio Borio (2017, p. 2), Chief Economist at the Bank for International Settlements, was quite candid: 'the response of inflation to a measure of labour market slack has tended to decline and become statistically indistinguishable from zero. In other words, inflation no longer appears to be sufficiently responsive to tightness in labour markets.' In other words, the inverse relationship between unemployment and inflation has disappeared. This is a devastating conclusion. Without these two relationships, monetary policy has lost its essence, and it begs for a complete rethinking of what monetary policy is and what it does. To better understand this, it is to post-Keynesian economists that we must turn our attention.

Monetary policy from a post-Keynesian perspective I

The post-Keynesian (and heterodox) perspective on central banking and monetary policy-making is widely different from the mainstream view. The differences stem from a series of different conceptions at both the positive and the normative levels. On the one hand, the nature and role of money and interest rates are essentially different from the mainstream perspective. On the other hand, monetary policy-making has an important influence on both income and wealth distribution according to the heterodox perspective, but its actual impact on the price level should not be overestimated. Hence, interest rate policies should consider their effects on income and wealth distribution, rather than focusing on the stability

Figure 7.1 The triangular nature of money

of the level of consumer prices, which they cannot influence as claimed by the mainstream.

The nature of money

First, the nature of money, hence also of central bank money, cannot be reduced to a particular commodity or financial asset, whose supply and demand depend on the 'market mechanism' leading to an 'equilibrium' price (or rate of interest, in the case of money). Money is not a commodity or a financial asset, but the means of final payment in the sense that it allows agents to finally pay (extinguish) their debts on any markets. Each bank in fact issues money in a triangular operation that involves the payer and the payee, as a result of which the latter has no further claim on the former, since they obtain the right to dispose of a bank deposit.

So assume you purchase (consume) a good from a store; you are now in debt vis-à-vis the store, which you extinguish by using cash (bank money), or your bank account. In this sense, your payment is finally settled through the bank (Figure 7.1).

Contrary to the conventional wisdom, however, the bank deposit does not stem from pre-existing saving, the origin of which would remain mysterious (as it cannot be explained logically). All bank deposits result, in fact, from a bank loan, and this is why a central bank must exist, to avoid banks moving 'forward in step' (Keynes, 1930/1971, p. 23) without any endogenous limit to money creation, because this would be a major factor of financial crises (Box 7.9 and Figure 7.2; see Chapter 4 on the monetary circuit). Indeed, the existence of a central bank closely depends on the book-entry nature of money (see Rochon and Rossi, 2013). Its role is to make sure that all payments on the interbank market are final rather than just promised, as would

BOX 7.9

BOOK-ENTRY NATURE OF MONEY

Mainstream theory assumes that money is exogenous and can be dropped from a helicopter and imposed on the economic system. Post-Keynesians, however, see money as the result of a prior loan and as endogenous to the needs of the economic system: the supply of money depends on the borrowing needs of private sector agents, and is closely tied to their production and investment needs. When a bank agrees to grant a loan, provided the borrower is deemed creditworthy, the loan is registered on the assets side of the bank's balance sheet, while a deposit simultaneously appears on the liabilities side of the same balance sheet (Figure 7.2).

Figure 7.2 The balance sheet of a commercial bank

Assets	Liabilities
Loan	Deposit

occur if banks would be allowed to pay simply by issuing their own acknowledgement of debt.

The triangular nature of money, as described above between a household and a firm using bank money, also applies to banks in their relationship with the central bank. To be sure, Figure 7.1 assumed that the payment was made between a household and a firm using the same bank, whereas in reality two different banks are most likely involved. So, when a consumer purchases a good, there are in fact two debts being created: the debt between the household and the firm, which is extinguished when the payment is settled through the bank, and a second debt involving the bank of the household and the bank of the firm. This debt, unlike the other, is settled only using central bank money (reserves), as shown in Figure 7.3.

As a result of the book-entry nature of money, the central bank intervenes always and everywhere when there is a payment to be settled between any two particular banks. It is therefore a matter of routine for a central bank to issue its own acknowledgement of debt in order for a payer bank to settle its debt against any other bank on the interbank market. This final payment at the interbank level may also involve the central bank as a financial intermediary, in the sense that the latter provides both money and credit to the paying bank. In particular, the central bank can lend to any bank participating in a domestic payments system the funds that this bank needs in order to settle

Figure 7.3 The triangular nature of money with a central bank

its debt to a third party. This may occur against collateral (that is, a series of financial assets that the paying bank provides to the central bank as a guarantee of repayment), or it might also occur without collateral (in an emergency situation, particularly when the paying bank is 'too big to fail' without provoking a number of negative outcomes across the domestic or global financial system). When the central bank intervenes in this regard, it plays the role of a lender of last resort, as nobody else is willing to lend to the bank in need of funding its own payments. The central bank may charge a rate of interest to the borrowing bank, which is usually higher than the interbank market rate of interest, as a penalty for the bank not being able to raise funds on that market. This suffices to induce banks to manage their own business in order to limit the number of instances where they need to turn to the central bank as a lender of last resort.

The intervention of the central bank as a lender of last resort confirms that the amount of central bank money is not set by the monetary authority (exogenously), but actually depends on the needs of the set of banks participating in the domestic payments system (endogenously). The 'helicopter view' of money, hence the money multiplier, therefore cannot account for the working of a monetary economy, thus invalidating the orthodox perspective on factual grounds. The rate of interest set by the central bank is indeed an exogenous magnitude, and does not depend on the market rates of interest; rather, it influences them (although not by a one-to-one relationship and with variable time lags, also depending on the size, type, and private or public ownership of each bank participating in the domestic payments system).

This empirical evidence against the orthodox perspective is also a critique of the alleged influence that monetary policy can exert on the general price level, hence the doubt about the capacity of a central bank to guarantee price

stability through its own rates of interest, as explained above. The orthodox argument that an interest rate hike reduces demand on the market for produced goods and services, thereby lowering inflationary pressures, can be questioned with regard to the likely impact of monetary policy tightening on banks' own rates of interest. If firms have to pay higher rates of interest because the central bank has raised its own interest rates for banks borrowing from it, then the firms' output will be sold at a higher price (hence an increase in the price level) owing to higher production costs due to bank lending.

The income distributive nature of monetary policy

This chapter has argued that monetary policy has failed in its intended purpose. The underlying theory regarding monetary policy and its transmission mechanism simply do not find empirical support, with the result that the conduct of monetary policy must now rely on luck rather than theory. This is not to say that monetary policy has no role to play in impacting economic activity, but rather, the transmission mechanism from changes in the rate of interest to economic activity is different than what mainstream economists believe. Indeed, rather than impacting on the price level as claimed by orthodox economists, post-Keynesians argue that a central bank's interest rate policy can actually affect income and wealth distribution, making it more concentrated at the top of the relevant pyramid.

There are two ways in which monetary policy can affect income distribution: through (1) the income channel; and (2) the wealth channel, as depicted in Figure 7.4.

The income channel can be divided further into a direct and an indirect mechanism, the first of which concerns the impact of changes in the rate of interest directly on the income of individuals: rentiers and workers. This works through the bond market and on bond holders, or rentiers as we call them: individuals whose income arises not from work, but from simply owning a class of assets. When interest rates increase, say the rate of interest on a ten-year bond, the higher rate translates into a higher return for bond holders. In this sense, lower rates will tend to reduce the income flow to rentiers. In this context, the rate of interest itself is an income distributive variable; a central conclusion of post-Keynesian economics.

The second, albeit indirect, mechanism works through the labour market. Changes in the rate of interest may have effects on labour markets, unemployment, and thus the income of workers. For instance, as the rate of interest falls in a recession, this may encourage the hiring of workers, a drop in

Figure 7.4 The income distributive effects of monetary policy

unemployment, and thus an increase in total wages, not to mention that as unemployment falls, workers may also be able to demand higher wages.

As for the wealth channel, consider, for instance, when policy rates of interest diminish, as in the aftermath of the global financial crisis that erupted in 2007–08, with the aim of supporting economic activity and employment levels. In fact, as already pointed out, neither firms nor households will be induced to increase their borrowing from the banking sector if they fear being unable to repay their debt (and the relevant interest) when it matures. Rather, this reduction of interest rates will spur financial transactions, thus inflating an asset price bubble that is further reinforced by the so-called 'wealth effect', which consists in feeling richer when one's assets are priced higher on the relevant market. Clearly, wealthy individuals whose assets are priced higher as a result of a reduction in policy rates of interest will not increase their spending to buy a series of consumption goods, thereby supporting economic activity, but rather increase their spending on real estate and financial markets, thus increasing the relevant asset prices. Holders of these assets will thereby feel richer, giving rise to an upward spiral that could inflate a bubble, threatening the financial stability of the whole system.

The conclusion to be drawn from this short analysis is that monetary policy may work first and foremost through income and wealth distribution. Incremental changes in interest rates affect the distribution of both income and wealth between rentiers and workers, and between households. Since we know that poorer individuals spend a greater proportion of their income than wealthier ones, a policy that redistributes income toward workers is better. In this sense, a permanent policy of low interest rates is to be considered. Indeed, this is what Rochon and Setterfield (2007, 2008, 2012) have advocated in their work on post-Keynesian interest rate rules. Of course, this policy would then require governments to adopt a proper regulatory framework to prevent financial bubbles.

Yet, the past three decades have shown the fragility of our economic system when monetary policy is not accommodative to workers. Lavoie and Seccareccia (1999) and Seccareccia and Lavoie (2016) showed that monetary policy has consistently favoured rentiers, thereby exacerbating an already unequal distribution of income and wealth. In other words, monetary policy has acted as an incomes policy that protected rentiers.

These phenomena, in addition to what we discussed above, are also relevant to criticizing the inflation-targeting strategy that has been in fashion at central bank level and within orthodox circles since the early 1990s, when an increasing number of monetary authorities abandoned monetary targeting in favour of adopting a policy strategy based on an explicit target for the rate of inflation, thereby considered as the principal (if not unique) goal of monetary policy until the financial crisis erupted in 2007–08. Indeed, inflation-targeting strategies and their apparent success in reducing and then keeping inflation rates at a low and stable level across a variety of countries since the early 1990s have not been in a position to avoid a major crisis, in the aftermath of the financial and real estate bubble that burst in the United States in 2006–07. Quite the contrary: these strategies could have been a factor of the bubble, and the ensuing crisis, as they led a number of central banks – first and foremost the US Federal Reserve – to keep their policy rates of interest too low for too long, thereby inducing several banks as well as non-bank financial institutions, in the absence of sound macroprudential regulations, to profit from this policy stance in order to increase both their lending volumes and their profits in an unsustainable way for the whole economic system.

The problem in this regard concerns the theoretical framework supporting inflation-targeting strategies. As stated above, based on monetarism, this framework considers that a monetary authority should guarantee price

stability and nothing else, because this provides the best macroeconomic environment for other categories of agents to contribute to economic growth and maximum employment levels. In other words, this approach argues that stable prices and low inflation are the best economic outcome possible.

The role of rationality and the independence of central banks

The above explanation argues that agents (firms and households) are equipped with so-called 'rational' expectations. Now, this is a very important assumption: in economics, rational expectations imply that agents possess all the required information and are able to calculate the most optimal solution, every time. Moreover, this solution is assumed always correct, since agents are thought of as knowing with certainly what the 'correct model' is, and as such treat all the newly available information in the right way. This suggests that the impact of changes in the rate of interest is fully anticipated and as such has very minor consequences in the long run. In this context, only non-anticipated changes in the rate of interest aimed at surprising agents can have an impact (albeit short-lived) on inflation that should support economic growth via a reduction of real wages and real interest rates (both of which induce firms to increase investment and employment levels, in the view of orthodox economists).

These arguments in fact have been instrumental in supporting central bank independence from any government pressures, arguing that these pressures (for instance, to 'monetize' public debt when a government's finance minister is unable to balance the public sector's budget through tax receipts or financial markets) eventually lead to overshooting the inflation target, without any positive influence on so-called 'real magnitudes' (such as economic growth, labour productivity and employment levels). As stated above, according to the mainstream view, agents are rational and have rational expectations, so they cannot be led astray by surprise inflation, as in that case they anticipate a rate of inflation higher than the central bank's publicly announced target and thus behave 'as if' the rate of inflation were already higher than the one officially targeted by the central bank (which therefore suffers from being dependent on the government, as it does not meet its own objective).

This argument has been further reinforced by referring to the so-called 'sacrifice ratio', that is, the ratio measuring output and employment losses as a result of a reduction in the rate of inflation. Mainstream literature points out that these losses are lower when the central bank is independent from the government, arguing that a central bank's independence enhances monetary policy credibility, which is instrumental in making sure that an inflation-

targeting strategy elicits no output and employment losses, as shown by expression (7.5):

$$Y_t = Y^* + \beta\,(\pi_t - \pi^e) \tag{7.5}$$

where Y_t is current output, Y^* is full employment (or potential) output, β is a positive parameter, π_t is the current rate of inflation, and π^e is the rate of inflation expected by economic agents. A central bank's credibility would be instrumental in order to make sure that inflation is on target without generating output and employment losses (in that case $Y_t = Y^*$). As a result, a restrictive monetary policy stance aimed at reducing the measured rate of inflation would give rise to no output and employment losses, if the central bank is independent of the government, as this is enough to make sure that both the central bank and its policy stance are credible.

Now, the problem with this reasoning is that agents are not 'always and everywhere' rational, as they do not have all the information and the knowledge required to be rational. Rather, as Keynes pointed out, the future is unknown and unknowable. Further, there are too many variables and reciprocal influences among their set to be able to appraise the actual working of the whole economic system with a particular model or series of models. Hence, the 'pretense-of-knowledge syndrome' (Caballero, 2010) that affects mainstream economics and the ensuing monetary policy-making is a dangerous factor of financial instability and economic crisis, as it gives a false perception of security in simply targeting price stability on the goods market by contemporary central banks. In fact, as heterodox economists have pointed out, inflation-targeting central banks have been inflicting an anti-growth bias to their economic systems, without preserving the latter from financial bubbles that gave rise to system-wide crises once they burst. A clear example is the monetary policy stance of the European Central Bank. Before the euro area crisis erupted towards the end of 2009, the bank's inflation-targeting strategy inflated a credit bubble in a variety of so-called peripheral countries across the euro area, while it hindered economic activities as the monetary policy rates of interest were not reduced until measured inflation had fallen below 1 per cent, but increased as soon as expected inflation (a virtual magnitude) was close to 2 per cent, even though measured inflation was much lower than that (see Chapter 14 for analytical elaboration on this).

Monetary policy from a post-Keynesian perspective II

As noticed above, post-Keynesians and heterodox economists argue in favour of a very different monetary policy stance: post-Keynesians rank the

well-being of citizens above the need to contain inflation. In their view, this leads to the policy conclusion that central banks should contribute to output stabilization and employment maximization, as well as sustainable economic growth, while also considering the impact of interest rate policy on income and wealth distribution across the economic system, without neglecting the fact that a central bank should contribute to financial stability for the system as a whole. This puts the conventional policy tools used by central banks (namely, interest rates, open-market operations, repurchase agreements and reserve requirements) under stress, as they are not appropriate, individually and as a whole, to ensure that the above set of policy goals are fulfilled adequately.

Let us revisit briefly the New Consensus Macroeconomics discussed earlier (equation 7.4), and specifically the interest rate rule, which we wrote as follows:

$$i_t = i_{t-1} + \alpha\,(\mu_t - \mu^*) \tag{7.4}$$

We can generalize this rule (many economists write it in different ways), and end up with a specific rule called the 'Taylor rule' (see Taylor, 1999):

$$i_t = r + \pi_t + \alpha\,(\pi_t - \pi^*) + \beta\,[(Y_t - Y^*)/Y^*] \tag{7.6}$$

where r is the so-called 'natural' (or equilibrium) rate of interest, π^* is the target rate of inflation, and α and β are positive parameters reflecting the importance of inflation and output gaps respectively in the central bank policy reaction function, that is, the difference between observed and targeted inflation, and between observed and potential output, respectively (see Box 7.9).

This policy rule has been criticized on several grounds. First, it integrates neither exchange rate issues nor the problems stemming from financial instability, perhaps because it is difficult to determine exchange rate misalignments and to define financial instability. Second, the notion of a 'natural' rate of interest is a figment of the imagination (Box 7.10), as it stems from a dichotomous view of the working of an economic system that is conceived as made up of a 'real' sector and a 'monetary' sector that would be separated from each other.

In the 'real' sector, equilibrium would be attained when savings are equal to the amount of firms' desired investment, thereby determining the 'natural' interest rate. In the monetary sector, by contrast, money supply and money

BOX 7.10

THE NATURAL RATE OF INTEREST

In economics, the natural or neutral rate of interest is the rate at which savings equal investment, and is simultaneously where inflation is equal to the target rate, and the economy is growing at its optimal level. It is a core principle in the conduct of monetary policy. The only problem is that it cannot be calculated. According to Claudio Borio, the Chief Economist at the Bank for International Settlements, 'the natural rate is an abstract, unobservable, model-dependent concept' (Borio, 2017, p. 8). This is yet more evidence of the problems associated with mainstream economic thinking.

demand (two separate and independent forces in the orthodox view) would determine, at equilibrium, the market rate of interest; which may thus differ from the natural rate, thereby generating an inflationary or deflationary pressure across the whole economic system.

Third, the 'Taylor rule' ignores the fact that interest rate policies affect both income and wealth distribution across this system, which is a major issue for financial stability as well as for macroeconomic stabilization, as explained above. In light of all these critiques, the heterodox approach to interest rate policy offers two alternative views: one that has the merit of including macroeconomic stabilization in a central bank's objectives, and another that focuses on the distributional impact of changing the rates of interest set by the monetary authority, as discussed above. In the former view, central banks are important actors in influencing the economic performance of the relevant countries, as monetary policy is in a position to support or hinder economic growth not just through its impact on price stability, but also via output and employment stabilization. In the latter view, by way of contrast, monetary policy impacts upon income and wealth distribution via the setting of interest rates, which tends to favour the owners of financial capital (the rentiers) in that the targeted rate of inflation is lower (and the policy rates of interest are higher) than what is needed in order to reduce the income share of rentiers and to increase the income share of wage earners with a view to inducing economic growth in a sustainable way for the whole system.

Conclusion

This chapter aimed at exploring and contrasting various approaches to monetary policy. We argued that the mainstream approach, based on the Taylor

rule and inflation targeting, finds very little empirical support for its theory, especially given the collapse of the Phillips curve relationship. This suggests that its fine-tuning approach of incremental changes in interest rates must be discarded. In its place, post-Keynesians correctly argue that monetary policy works mainly through income distribution, both directly and indirectly, through labour markets.

Of course, mainstream economists have not accepted this reality, and instead continue to rely on counter-cyclical changes in interest rates. Yet, for reasons explained in this chapter, this approach does not work, and central banks are forced to increase interest rates several times in an elusive quest to reach a fictional natural rate of interest. In doing so, interest rates are pushed so high that in the end they have a negative impact on economic activity, proving that there is nothing neutral on the road to finding the natural rate of interest. In the end, monetary policy collapses the economy, and reveals itself at best as being a blunt tool that causes tremendous harm.

The above discussion shows that monetary policy cannot be considered merely a technical matter that should be left to 'technicians', or perhaps to an automatic pilot, without any political economy consideration. This echoes the famous argument already raised in the 1930s by Ralph George Hawtrey, who wrote that monetary policy is an art rather than a science (Hawtrey, 1932, p. vi). As such, and in light of their socioeconomic consequences, monetary policy decisions should be taken as a result of a systemic appraisal of their effects and considering the common good, that is, the well-being of the whole population affected by them.

REFERENCES

Bofinger, P. (2001), *Monetary Policy: Goals, Institutions, Strategies, and Instruments*, Oxford: Oxford University Press.

Borio, C. (2017), 'Through the looking glass', lecture delivered at OMFIF, 22 September, London.

Caballero, R.J. (2010), 'Macroeconomics after the crisis: time to deal with the pretense-of-knowledge syndrome', *Journal of Economic Perspectives*, **24** (4), 85–102.

Cynamon, B., S. Fazzari and M. Setterfield (2013), 'Understanding the Great Recession', in B. Cynamon, S. Fazzari and M. Setterfield (eds), *After the Great Recession: The Struggle for Economic Recovery and Growth*, Cambridge, UK: Cambridge University Press, pp. 3–30.

Eichner, A. (1976), *The Megacorp and Oligopoly*, Cambridge: Cambridge University Press.

Eichner, A. (ed.) (1979), *A Guide to Post-Keynesian Economics*, Armonk, NY: M.E. Sharpe.

Eichner, A. (1985), *Toward a New Economics: Essays in Post-Keynesian and Institutionalist Theory*, Armonk, NY: M.E. Sharpe.

Eichner, A. (1987), *The Macrodynamics of Advanced Market Economies*, Armonk, NY: M.E. Sharpe.

Friedman, M. (1969), *The Optimum Quantity of Money and Other Essays*, Chicago, IL: Aldine Publishing.

Friedman, M. (1987), 'Quantity theory of money', in J. Eatwell, M. Milgate and P. Newman (eds),

The New Palgrave: A Dictionary of Economics, Vol. IV, London and Basingstoke: Macmillan, pp. 3–20.

Friedman, M. and A. Schwartz (1963), *Monetary History of the United States 1867–1960*, Princeton, NJ: Princeton University Press.

Hawtrey, R.G. (1932), *The Art of Central Banking*, London: Longmans, Green & Company.

Holt, R. and S. Pressman (2011), *A New Guide to Post Keynesian Economics*, London: Routledge.

Keynes, J.M. (1930/1971), *A Treatise on Money, Volume 1: The Pure Theory of Money*, reprinted in *The Collected Writings of John Maynard Keynes, Volume V*, London and Basingstoke: Macmillan.

Keynes, J.M. (1979), *The Collected Writings of John Maynard Keynes, Volume XXIX: The General Theory: A Supplement*, London and Basingstoke: Macmillan.

Krugman, P. (2018), 'Why was Trump's tax cut a fizzle?', *New York Times* blog, 18 November, available at https://www.nytimes.com/2018/11/15/opinion/tax-cut-fail-trump.html.

Lavoie, M., L.-P. Rochon and M. Seccareccia (2010), *Money and Macrodynamics: Alfred Eichner and Post-Keynesian Economics*, Armonk, NY: M.E. Sharpe.

Lavoie, M. and M. Seccareccia (1999), 'Interest rate: fair', in P. O'Hara (ed.), *Encyclopedia of Political Economy*, London, UK and New York, USA: Routledge, pp. 543–5.

McLeay, M., A. Radia and R. Thomas (2014), 'Money creation in the modern economy', *Bank of England Quarterly Bulletin*, **54** (1), 14–27.

Phillips, A.W. (1958), 'The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957', *Economica*, **25** (100), 283–99.

Rochon, L.-P. and S. Rossi (2013), 'Endogenous money: the evolutionary versus revolutionary views', *Review of Keynesian Economics*, **1** (2), 210–29.

Rochon, L.-P. and M. Setterfield (2007), 'Interest rates, income distribution and monetary policy dominance: post Keynesians and the "fair rate" of interest', *Journal of Post Keynesian Economics*, **30** (1), 13–41.

Rochon, L.-P. and M. Setterfield (2008), 'The political economy of interest rate setting, inflation, and income distribution', *International Journal of Political Economy*, **37** (2), 2–25.

Rochon, L.-P. and M. Setterfield (2012), 'A Kaleckian model of growth and distribution with conflict-inflation and Post-Keynesian nominal interest rate rules', *Journal of Post Keynesian Economics*, **34** (3), 497–520.

Seccareccia, M. and M. Lavoie (2016), 'Income distribution, rentiers and their role in a capitalist economy: a Keynes–Pasinetti perspective', *International Journal of Political Economy*, **45** (3), 200–23.

Sharpe, S. and G. Suarez (2014), 'Why isn't investment more sensitive to interest rates: evidence from surveys', Federal Reserve Board Finance and Economics Discussion Series, No. 2014-002.

Taylor, J.B. (1993), 'Discretion versus policy rules in practice', *Carnegie-Rochester Conference Series on Public Policy*, **39**, 195–214.

Taylor, J.B. (1999), 'Monetary policy guidelines for employment and inflation stability', in R.M. Solow and J.B. Taylor (eds), *Inflation, Unemployment, and Monetary Policy*, Cambridge, MA: MIT Press, pp. 29–54.

A PORTRAIT OF ALFRED S. EICHNER (1937–88)

Born in the United States on 23 March 1937, Alfred Eichner was a leading, though considerably undervalued, post-Keynesian economist who contributed to the advancement of pricing theory, the theory of economic growth and income distribution, and monetary theory and policy. He is considered by many as one of the founders of the US post-Keynesian school.

Eichner received his doctorate from Columbia University in New York, and taught at Columbia University; Purchase College, State University of New York (SUNY); and at Rutgers University, where he was lecturing at the time of his death on 10 February 1988.

Eichner has written numerous articles and books, but is perhaps best remembered for two books in particular: his *Guide to Post-Keynesian Economics*, published in 1979, which offered for the first time a complete and concise introduction to various themes within post-Keynesian economics, and *The Macrodynamics of Advanced Market Economies*, published in 1987, which contains certainly his most important insights into the workings of an advanced market economy.

In his *Guide*, which contains a foreword by Joan Robinson, Eichner brought together, for the first time, ten leading post-Keynesian economists of the time to discuss topics of great relevance to heterodox and critical economics, including a chapter on 'Monetary factors' by Basil Moore. This marked an important first step in the institutionalization of post-Keynesian economics, and coincided with the creation of the *Journal of Post Keynesian Economics,* in the autumn of 1978. In 2011, Richard Holt and Steve Pressman edited *A New Guide to Post Keynesian Economics* (strangely enough, without a hyphen).

In his *Macrodynamics* book, Eichner discusses important themes of pricing, economic growth and distribution. He rejected the neoclassical price theory of supply and demand, and argued instead that prices were determined by a mark-up over the costs of production, thereby bringing together the importance of social classes. Eichner also discusses issues touching on central banking and monetary policy that were at the time, and still are today, considerably innovative and ahead of his time. Eichner showed that he had an incredible grasp of monetary details that are as relevant today as they were at the time (see Lavoie et al., 2010).

Eichner was not only an avid author and professor, but also somewhat of an activist, having testified numerous times on Capitol Hill before several Congressional and other legislative committees.

Among his other books are *The Megacorp and Oligopoly* (1976), which explored price setting in an oligopolistic setting and how this influences economic growth and economic stability, and *Toward a New Economics: Essays in Post-Keynesian and Institutionalist Theory* (1985), in which Eichner further develops his ideas on the 'megacorp', an approach to labour he calls anthropogenic, modelling and post-Keynesian economics, empirical research, and views on social democracy.

EXAM QUESTIONS

True or false questions

1. The Taylor rule is used to strengthen fiscal spending rules.
2. Monetary policy works through the income distributive channel.
3. Inflation targeting regimes have been very successful.
4. The natural rate of interest in central to the mainstream approach to monetary policy.
5. Central banks are powerful institutions that intervene on a regular basis as a lender of last resort.
6. Central banks are always able to set the rate of interest.
7. Monetarism as an economic theory was a failure.
8. Commercial banks always need prior reserves before making loans to the private sector.
9. Because of the book-entry nature of money, the central bank intervenes always and everywhere when there is a payment to be settled between any two particular banks.
10. Fine-tuning monetary policy is an efficient use of monetary policy.

Multiple choice questions

1. The Phillips curve illustrates:
   a) the relationship between the money supply and unemployment;
   b) the relationship between changes in unemployment rates and the level of economic activity;
   c) the relationship between changes in the price level and unemployment;
   d) the impact of money supply growth on inflation.
2. The Taylor rule is:
   a) a rule about how much the money supply should be increasing each year;
   b) an interest rate rule dependent on a target rate of inflation;
   c) a measure of unemployment at less than full-employment equilibrium;
   d) a ratio between fiscal deficits and gross domestic product (GDP) growth.
3. The role of lender of last resort suggests:
   a) the central bank stands ready to lend to banks if they need liquidity;
   b) the government stands ready to inject spending into the economy;
   c) the firms are deemed less creditworthy;
   d) bank lending to the private sector under proper rules of profit maximization.
4. Changes in the rate of interest affect income distribution through:
   a) the income channel and the ratio of imports to GDP;
   b) the income channel and the wealth channel;
   c) the wealth channel and the lending channel;
   d) the elimination of the effect on imports.
5. The concept of rational expectations explains:
   a) all agents possess all required information and are thus able to calculate the most optimal solution, every time;
   b) when expectations about the rate of inflation coincide with expectations about the rate of unemployment;
   c) when agents act rationally in deciding whether to increase consumption;
   d) agents have less than optimal information, but can still choose optimal solutions when the economy is growing at full employment.
6. The sacrifice ratio is:

a) the loss in consumption when unemployment remains the same;
b) the personal sacrifices made by consumers in saving for the future;
c) the decrease in inflation when unemployment increases;
d) losses in output and employment as a result of a reduction in the rate of inflation.

7. The natural rate of interest is:
a) a useful guide to monetary policy;
b) equal to the growth rate of GDP;
c) impossible to calculate and therefore not relevant to monetary policy;
d) always equal to the rate of interest on government deposits.

8. The quantity theory of money:
a) considers the relationship between money and its velocity of circulation, on the one side, and the price of output, on the other side;
b) measures the number of times goods circulate in the economy;
c) examines the triangular relationship of money;
d) measures the value of money and how it grows through policy.

9. The Phillips curve today:
a) remains as relevant as it did three decades ago;
b) is more relevant today for monetary policy than before;
c) has flattened, making it irrelevant for monetary policy today;
d) has transformed into a super Phillips curve.

10. According to Milton Friedman:
a) inflation is the result of too many goods being produced;
b) inflation is always and everywhere a monetary phenomenon;
c) inflation is always and everywhere an income distribution phenomenon;
d) inflation is the result of diminishing fiscal spending.

Part III

The macroeconomics of the short and long run

8

Theories of consumption

Stavros A. Drakopoulos

OVERVIEW

This chapter:

- explains the role of consumption expenditures in modern economies and their significance for the determination of the level of output and employment in an economy;
- presents the theory of intertemporal choice that forms the basis of mainstream consumption functions;
- discusses Keynes's approach to consumption, and particularly his criticism of the standard model of consumer behaviour, his emphasis on the role of consumption for the level of employment, and his analysis of aggregate consumption patterns;
- describes the main mainstream theories of consumption, which are the life-cycle income hypothesis, the permanent income hypothesis and the random walk theory of consumption;
- explores the heterodox approaches to consumption, focusing mainly on the relative income hypothesis;
- shows the consequences of consumption theories for the effectiveness of economic policies towards unemployment and economic downturns.

KEYWORDS

- **Consumption expenditures:** Expenditures made by the household sector on final goods and services. Consumption expenditures are the major category of expenditures as a share of gross domestic product (GDP). The other three are private investment expenditures, government purchases and net exports.

- **Intertemporal choice:** The theory assumes that the consumer maximizes utility out of consumption subject to an intertemporal budget constraint, a measure of the total resources available for present and future consumption.
- **Fiscal policy and consumption:** The effectiveness of economic policies towards reducing unemployment depends on the role of current or relative income in the consumption function.
- **Keynesian consumption function:** It is an important theoretical concept used in Keynesian macroeconomics and shows the relationship between consumption and current disposable income.
- **Life-cycle income hypothesis:** It assumes that individuals maximize utility, which is expressed as a function of the individual's consumption stream over the span of their lifetime.
- **Permanent income hypothesis:** Permanent income is an individual's income over their lifetime. Consumers use their savings (or borrow) in an attempt to smooth consumption between good and bad years.
- **Relative income hypothesis:** The theory assumes that consumers will be influenced by the behaviour of other consumers' consumption and that consumption relations are not reversible in time.

Why are these topics important?

Consumption represents a large part of expenditures on product markets in modern economies. Its share with respect to GDP is around 70 per cent in most advanced economies and even more in less advanced ones. Consumption is important for aggregate demand – that is, the total demand for all goods and services in the economy – since, according to Keynesian theory, aggregate demand determines the level of output and employment in an economy: the more we demand, the more we produce, and the more we create employment. Also, income that is not consumed is saved, and savings have a large impact on the growth of an economy. Thus, consumption is important to understand savings, capital stock, investment, employment and income growth. But there is more: the effectiveness of economic policy is also closely related to the nature of the consumption function.

To understand this, we need to consider, as Keynes did, that consumption depends on current disposable income, that is, current income minus taxes. In turn, the marginal propensity to consume (to wit, the change in consumer spending due to a change in income) determines the magnitudes of government expenditure and tax multipliers. To put it simply: how much GDP increases or decreases when the government increases or decreases spending in the economy, depends on how much we spend on consumption. It also determines the magnitude of the private investment multiplier.

This suggests that the effectiveness of fiscal policy – that is, changes in taxation and government spending – to smooth out economic downturns, and thus large fluctuations of produced output and employment, is linked to the level of fiscal multipliers and therefore to the nature of the consumption function.

As a reaction against Keynesian economic policies, mainstream economics has sought to downplay the importance of fiscal policy, and greatly diminished its role. One way of doing this is by undermining the importance of consumption as a function of income, since if a large part of consumption is independent of income (that is, autonomous consumption), as in many mainstream consumption functions, then the fiscal multiplier is not very significant and the role of fiscal policy is greatly reduced (see Bunting, 1989). In other words, the increase or decrease of GDP due to government spending changes has little to do with how much we spend in consumption. This is the policy conclusion of mainstream consumption theories such as the life-cycle hypothesis and the permanent income hypothesis that will be explored in this chapter.

The mainstream view: intertemporal choice and consumption function

One of the core components of mainstream economic theory is the concept of economic rationality, with its most frequent expression in the model of the rational consumer. In the framework of mainstream economic theory, rational consumers are assumed to behave as selfish utility maximizers. More specifically, in consumer theory rational consumers obey the axioms of rational choice and maximize their utility function subject to the budget constraint. Consequently, individual demand curves are derived from this standard model of utility maximization. The same model is then extended to include intertemporal choice – that is, the choice between consumption today versus consumption at some point in the future – which forms the basis of mainstream consumption function. In the same way the theory of rational consumer is used to derive individual demand functions for goods and services, intertemporal maximization is used to derive aggregate consumption functions. The neoclassical economist Irving Fisher set the theoretical basis of this approach in his *Theory of Interest* (Fisher, 1930). In this framework it is assumed that consumers are rational, forward-looking agents and choose consumption levels for the present and future so as to maximize lifetime satisfaction. Consumers' choices are subject to an intertemporal budget constraint, a measure of the total resources available for present and future consumption.

In a simple two-period model, there are two periods: period 1 (the present) and period 2 (the future). Let Y_1 and Y_2 be income in period 1 and 2, respectively. Let C_1 and C_2 be consumption in period 1 and period 2, respectively, and S is savings. The consumers' budget constraint in the first period is:

$$Y_1 = C_1 + S \text{ or } S = Y_1 - C_1 \tag{8.1}$$

If $S > 0$ the consumer is saving, and if $S < 0$ the consumer is borrowing. Saving (borrowing) yields (costs) an interest rate (r). In the next period, the budget constraint is therefore:

$$C_2 = Y_2 + (1+r)S \text{ or } C_2 = Y_2 + (1+r)(Y_1 - C_1) \tag{8.2}$$

The term $(1 + r)S$ is the rate of return of savings. We can rearrange terms to write:

$$(1+r)C_1 + C_2 = Y_2 + (1+r)Y_1 \tag{8.3}$$

If we divide the above expression by $(1+r)$, we get:

$$C_1 + \frac{C_2}{1 + r} = Y_1 + \frac{Y_2}{1 + r} \tag{8.4}$$

Equation (8.4) shows that the present value of lifetime consumption (on the left-hand side of the equation) equals the present value of lifetime income (on the right-hand side of the same equation). Present value means the value in terms of the consumption goods in period 1. In other words, $1/(1+r)$ is the relative price of future consumption in terms of current consumption: one unit of consumption today is equivalent to $1+r$ units of consumption tomorrow. Figure 8.1 shows the intertemporal budget constraint.

In Figure 8.1 the budget constraint shows all combinations of C_1 and C_2 that just exhaust the consumer's resources. At point E, consumption equals income in both periods. At points that are between B and E the consumer has positive savings, while at points that are between E and A the consumer has negative savings (that is, they borrow).

The budget constraint is combined with the notion of indifference curves, which show combinations of present and future consumption levels that leave the consumer equally happy. Moreover, higher indifference curves represent higher levels of happiness. As a utility maximizer, the consumer aims to be at the highest possible indifference curve given the consumer's budget constraint (Figure 8.2).

Figure 8.1 The intertemporal budget constraint

Figure 8.2 Optimization of intertemporal choice

In Figure 8.2 the optimal point is E, where the slope of the budget line just touches the highest possible indifference curve (I_2). Only at this point does the consumer maximize their utility out of consumption given a budget constraint. At point E, the slope of the indifference curve, which is the marginal rate of substitution (MRS) between present and future consumption, equals the slope of the intertemporal budget constraint ($1+r$). At point E, therefore, MRS = $1+r$.

Point E can change if there are changes in income (the budget line shifts) or changes in the rate of interest (the slope of the budget line changes). Intertemporal consumption theory indicates that consumption decisions depend only on the present value of lifetime income (as opposed to disposable income today). Since the consumer can borrow or lend between periods, the level of interest rate plays an important role in consumption decisions. The consumption function takes the general form:

$$C_{jt} = \mathrm{f}\,(PW_{jt}) \tag{8.5}$$

where C_{jt} denotes the consumption expenditures of individual j at time t, and PW_{jt} is the individual's corresponding present worth.

Fisher's intertemporal choice, as described above, is the conceptual basis of modern mainstream consumption theories today (such as the life-cycle hypothesis, and the random walk consumption function). But it is not without its critics. For instance, heterodox economics has directed a large amount of criticism toward this mainstream economic theory, largely based around the concept of rationality (sometimes referred to as '*Homo oeconomicus*') and its implications. One of the most important criticisms originated from the work of John Maynard Keynes, to which I now turn.

Keynes's approach to consumption

This section will discuss the original work of Keynes, and how it relates to consumption theory. Through this discussion, the limitations of mainstream theory will become clear, and its overall irrelevance for economic policy will be pointed out clearly.

There are three basic elements in Keynes's treatment of consumption: his criticism of the standard model of consumer behaviour, his emphasis on the role of consumption for the level of employment, and his analysis of aggregate consumption pattern. These will be discussed in turn.

Keynes's criticism of mainstream theory

Keynes's criticism of the mainstream theory of consumption can be divided into three arguments. Taken together, they show how Keynes was not very interested in explicitly formulating a consumer theory in the form that we find in contemporary microeconomics.

First, consider that contemporary microeconomics, based on the rational consumer and the marginalist theory of the economic agent (see Chapter 3), can be traced to the writings of the nineteenth-century founder of utilitarianism, Jeremy Bentham. Yet, there are strong indications that Keynes rejected the utility maximizing model (atemporal or intertemporal). His rejection of Benthamite hedonism and his ideas on probability and uncertainty clearly imply his distance from the standard model.

Bentham's 'calculus of pleasure and pain' was the basis of the marginalist theory of consumer behaviour, which can be found in leading nineteenth-century marginalist authors such as William Stanley Jevons, Léon Walras and Francis Ysidro Edgeworth (for a review and further discussion, see Drakopoulos, 1990). Keynes expressed serious doubts concerning the mainstream tradition originating from Bentham, especially in the form of (expected) utility maximization. In a very important passage written in 1939, Keynes considers himself and his circle as 'the first of our generation, perhaps alone amongst our generation, to escape from the Benthamite tradition' (Keynes, 1978b, p. 445). He continues by pointing out that '[i]t can be no part of this memoir for me to try to explain why it was such a big advantage for us to have escaped from the Benthamite tradition' (ibid., p. 445).

Second, Keynes's revolutionary conception of probability and uncertainty is another important point that puts him at odds with standard utility theory. According to Keynes, probabilities are either numerically indeterminate or undefinable (Keynes, 1978a, pp. 8–9; see also Lawson, 1988, pp. 42–4). Moreover, he thought a probability concept based on frequency (that is, repeated trials or repeated sampling from a population) is 'a wrong philosophical interpretation of probability' (Keynes, 1978a, p. 342). This is in sharp contrast with the mainstream expected utility approach, which views probability as based on frequency and as numerically measurable (see, for instance, Savage, 1962). For example, if a student scored an 'A' seven times out of the ten tests, then 7 is the frequency of scoring an 'A' and 7/10 * 100% = 70% is the relative frequency of scoring an 'A' in exams.

Third, related to this is Keynes's conception of uncertainty, which corresponds to a situation of numerically immeasurable probability. In other words, the uncertain knowledge of the future implies that 'we simply do not know' and cannot know events in the future. Accordingly, Keynes believed that uncertainty cannot be reduced, mainly because a numerical probability distribution is not known (events are not replicable). Keynesian uncertainty is thus radically different from the reducible and calculable uncertainty that is used in the expected utility model (see Dow, 1985, p. 156; Lawson, 1988,

pp. 46–52). In Keynes's (1936, p. 161) own words: '[It is a] characteristic of human nature that a large proportion of our purposive activities depend on spontaneous optimism rather than on mathematical expectation, whether moral or hedonistic or economic.' An example of Keynes's notion might be the following: the decision of an economic agent to invest does not depend so much on the expected future returns of their investment, but on spontaneity and inspiration.

The implication of Keynes's arguments is clear: in place of utility maximization, Keynes assigned extreme importance to psychological processes, which have nothing to do with economic calculus (see Dow and Hillard, 1995). He uses the term 'animal spirits' for these psychological processes. In an indicative statement, Keynes (1936, pp. 161–2) argues indeed that:

> Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits – of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities.

Keynes's serious reservations concerning the utility maximizing model lead him to an alternative formulation of consumption analysis, one that is not based on intertemporal analysis. Starting from the idea that consumption depends on both objective and subjective factors, Keynes states the following subjective factors, which he calls motives: enjoyment, short-sightedness, generosity, miscalculation, ostentation and extravagance. The corresponding list for saving behaviour is: precaution, foresight, calculation, improvement, independence, enterprise, pride and avarice (Keynes, 1936, p. 108). Keynes gives equal weight to these motives; something that is incompatible with the utility maximizing model. In general, Keynes approached the consumption decision from an entirely different angle than mainstream economists before him. His approach is essentially a psychological approach emphasizing also the sociological dimension of consumption (Drakopoulos, 1992; D'Orlando and Sanfilippo, 2010).

The significance of consumption in Keynes

Although other authors before Keynes had speculated about the relation between consumption and income, he was the first to make this relation explicit and to use it in a general model of the whole economy. The consumption function was an important component of Keynes's system and also a significant point of difference from the classical approach. In the general view

BOX 8.1

IRVING FISHER

Irving Fisher (1867–1947) was one of the earliest American neoclassical economists who effectively brought the ideas of the marginal school of economic thought to the United States. His first book, entitled *Mathematical Investigations in the Theory of Value and Prices* (1892/1961), was based on his doctoral thesis at Yale University. In this book, he made original contributions to utility theory and general equilibrium theory. In his subsequent works, he contributed to the quantity theory of money, which served as a basis for the school of economic thought known subsequently as 'monetarism'. Based on his analysis of monetary theory, he was a proponent of the full-reserve banking reform, where the general system of 100 per cent reserve backing for bank deposits is outlined. Fisher was also a pioneer of econometrics, including the development of index numbers. One of his major contributions was the theory of intertemporal choice in markets, which led him to develop a theory of capital and interest rates.

held by the classical economists before Keynes, the economy was always naturally tending towards full-employment equilibrium. It follows that, at least in the short run, income does not vary since its level is established at its full-employment level. Therefore, in the classical world variations in consumption and saving are not related to variations in income. Instead, the amount of income saved (and also the amount of income consumed) by an individual depends on the rate of interest, as explained above. Thus, the interest rate is the major determinant of the allocation of income between present and future (intertemporal) consumption (and savings). This idea is the basis of American economist Irving Fisher's (a contemporary of Keynes) intertemporal choice discussed above (Box 8.1). On the contrary, by rejecting the idea that the economy tends towards full employment, Keynes pointed out the differences between the actual level of income and the full-employment level. In Keynes's view, the level of aggregate demand determines equilibrium income. Given that consumption is a major part of aggregate demand (see Chapter 10), it was necessary to provide a theory of the behaviour of consumption expenditures.

Keynes's consumption function

In *The General Theory*, Keynes (1936) described consumption expenditures as an important component of national income. He further stated that current disposable income (after taxes) is the main determinant of consumption

expenditures. Keynes's approach to consumption has been called the absolute income hypothesis, because current disposable income only determines the level of consumption. With this as a basis, the Keynesian consumption function is usually written in a linear form:

$$C = \alpha + bY \tag{8.6}$$

where C is consumption expenditures, Y is current disposable income, α is consumption expenditures that are independent of income (or autonomous consumption), and b is the marginal propensity to consume (MPC) or the ratio of the change in consumer spending to a change in income. The MPC is also the slope of the consumption function. In short:

$$MPC = \frac{\partial C}{\partial Y} = b \tag{8.7}$$

The average propensity to consume (APC), which is the ratio of total consumption to total disposable income, is also given as:

$$APC = \frac{C}{Y} = \frac{a}{Y} + b$$

The Keynesian consumption function can be presented with a simple diagram, as Figure 8.3 shows.

Figure 8.3 The Keynesian consumption function

In Figure 8.3 consumption expenditures (C) are represented on the vertical axis while disposable income (Y) is on the horizontal axis. All the points of the 45° line show equality between consumption and disposable income. Consumption starts at α with slope b. At equilibrium point E all income is spent on consumption ($Y = C$), and savings are zero. For income levels lower than Y_1 consumption expenditures are higher than income, and therefore savings are negative (borrowing). For income levels higher than Y_1, consumption expenditures are lower than income and therefore savings are positive. For instance, for a high level of income Y_2, consumption is lower than income and savings are equal to the distance AB.

Keynes thought that the MPC is positive but less than one, and that the APC falls as income rises. In Keynes's (1936, p. 96) words:

> The fundamental psychological law, upon which we are entitled to depend with great confidence both *a priori* and from our knowledge of human nature and from the detailed facts of experience, is that men are disposed, as a rule and on the average, to increase their consumption, as their income increases, but not by as much as the increase in their income.

This implies that households with higher income will consume more (given that MPC > 0), will save more (given that MPC < 1), and that the APC will be falling as income increases. How the APC varies as income changes depends on a. In the normal case ($a > 0$), MPC < APC and households spend a decreasing share of their incomes as incomes rise. If $a = 0$, MPC = APC and spending is a constant proportion b of income.

It is important to emphasize once again that Keynes was not interested to ground his consumption function on the model of rational economic behaviour. Instead, he relied on aggregate psychological tendencies.

The economic policy implications of Keynes's approach are fairly well known. The magnitude of the MPC determines the magnitude of government expenditures and tax multipliers and thus the effectiveness of fiscal policy to maintain or restore full employment. In other words, the larger the MPC, the larger the multiplier effect (see Box 8.2 for proof). It also implies that as households spend a decreasing share of their incomes as society becomes richer, a greater proportion of investment will be required to maintain full-employment income levels.

In addition, and given that MPC < APC, a transfer of income from high-income groups to low-income consumers will raise the level of aggregate

BOX 8.2

THE IMPORTANCE OF THE MPC FOR THE MULTIPLIERS

In a simple closed-economy macroeconomic model, we have $Y = C + I + G$. National income (Y) comprises of consumption expenditures (C), private investment (I), and government expenditures (G). The equilibrium level of income (Y^*) is given by:

$$Y^* = \frac{1}{1-b}(a + I_0 + G_0 - bT_0)$$

where I_0, G_0 and T_0 are autonomous private investment, autonomous government expenditure and autonomous taxes, respectively. The ratio of the change in income to a change of investment is called the investment multiplier (m) and is given as $m = \frac{\Delta Y}{\Delta I}$.

In the simple Keynesian model, the private investment multiplier is:

$$\frac{dY^*}{dI} = \frac{1}{1-b}$$

Similarly, the government and the tax multipliers are:

$$\frac{dY^*}{dG} = \frac{1}{1-b}$$

$$\frac{dY^*}{dT} = \frac{-b}{1-b}$$

If, for instance, MPC = 0.8, the magnitude of private investment and government spending multiplier is 1/(1–0.8) = 5. This means that other things being equal, an increase of government spending by (say) $100 million will increase GDP by $500 million. Further, an increase in taxes by the same amount will diminish GDP by $400 million, given that the tax multiplier is –0.8/(1–0.8) = –4. If MPC = 0.9, then the magnitude of investment and government spending multiplier is 1/(1–0.9) = 10. Thus, the MPC determines the magnitude of fiscal multipliers (government and tax multipliers).

demand. This is also because the high-income groups have a lower MPC than low-income groups, given that MPC declines as income increases. The case for progressive taxation as an instrument of income redistribution is also justified from this argument.

Empirical testing of consumption functions

The first empirical estimations of Keynes's consumption function using aggregate time series data indicated a value of b (the MPC) around 0.75. More specifically, on the basis of annual United States (US) data for the period 1929–41, the estimated consumption function was reported being $C_t = 26.6 + 0.75Y_t$ (Ackley, 1960, p. 225). Other early studies were based on data capturing both income and consumption expenditures for some specified time period (cross-section). These budget studies indicated a non-proportional relationship between family income and family consumption. In other

BOX 8.3

SIMON KUZNETS

Simon Kuznets (1901–85) was a Russian-born economist who won the Nobel Prize in 1971 for his work on collecting, measuring and analysing economic statistical data. Kuznets earned a PhD in 1926 at Columbia University, USA, under the supervision of the well-known institutional economist Wesley Clair Mitchell. One of his main research interests was in economic fluctuations, where he set the basis for a systematic analysis of business cycles. His work provided empirical foundations to the Keynesian theory and played an important role in the dissemination of Keynesian policies after the Second World War. Kuznets tended to analyse the economy in connection and with the wider context of historical, demographic and social processes, an approach which is also attributed to his early research experience in his native country. In the framework of his studies on the formation of national income, Kuznets revealed long-term trends of output and income, consumption and savings, net investment, and so on. One important aspect of his work was the empirical testing of Keynes's absolute income hypothesis.

words, they showed that APC falls as income increases. They also pointed to a positive value of α. These findings confirmed the previous theoretical results that the MPC was less than the APC. In this framework, saving can be viewed as a 'luxury' good, whose share of overall income rises as people receive higher incomes (Venieris and Sebold, 1977, p. 365). Thus the pattern observed in cross-sectional consumption data meant that at a given point in time the rich in the population saved a higher fraction of their income than did the poor (see Koçkesen, 2008).

In the mid-1940s, American economist Simon Kuznets studied the characteristics of the consumption function based on his detailed reconstruction of US historical data on economic aggregates (Box 8.3).

Kuznets's work estimated aggregate consumption and income from 1869 to 1938. His results showed that the share of income consumed seemed to remain constant. Using overlapping decade averages of consumption and GDP, Kuznets (1946) showed that, except for the years of the Great Depression (1929–33), the APC in the United States over the period 1869–1938 fluctuated narrowly between 0.84 and 0.89. In other words, consumption and income tend to be proportionally related over a very long span of time series data. Although the two sets of empirical evidence are not necessarily inconsistent, Kuznets's (1946) findings suggested that the long-run behaviour of consumers might differ from their short-run consumption

Figure 8.4 Short-run and long-run consumption functions

patterns. More specifically, if MPC < APC as the ordinary least squares (OLS) estimates of the linear consumption function suggested, then the share of income consumed should have declined as income increased. Thus, short-run econometric studies found MPC < APC, while long-run data showed that MPC = APC (Figure 8.4). In essence, Kuznets's results suggested a consumption function of the form $C = kY$. This equation implies that MPC = APC = k. Further, the value of the MPC is much higher in Kuznets's function compared to Keynes's.

In Figure 8.4, the long-run consumption function (C_{LR}) has a slope equal to the long-run APC (and MPC). The short-run consumption functions (C_{SR1}, C_{SR2}) have a slope (MPC) that is smaller than the APC.

The apparently conflicting empirical evidence was the main reason for subsequent attempts towards a consumption theory that would provide reconciliation between the two sets of findings. One of the earliest theoretical approaches was the relative income hypothesis by James Duesenberry, which will be presented later. It also led to new theories of consumption, which were based on mainstream microeconomic foundations, such as the utility maximizing model with forward-looking agents. These theories were very critical of Keynes's psychological approach to consumption.

Mainstream consumption theories

Kuznets's empirical findings provided the initial stimulus for theoretical research on aggregate consumption patterns. The first theories to appear were the life-cycle hypothesis and the permanent income hypothesis. These two theories started by employing Fisher's intertemporal choice as the microeconomic foundation of aggregate consumption, thus rejecting Keynes's psychological-based approach to consumption. Current orthodoxy with the core assumptions of utility maximizing and forward-looking agents is essentially based on these two theories.

Life-cycle hypothesis

The life-cycle hypothesis of consumption function was developed mainly by Franco Modigliani (Box 8.4) and Richard Brumberg in 1954 (Modigliani and Brumberg, 1954).

Its underlying conceptual basis is that individuals maximize their utility of consumption over their life cycle, and not over their disposable income during, say, a year. In this sense, the basic tenet of the theory is the mainstream model of utility maximizing agents, which is based on the theory of rational consumer of mainstream microeconomics: rational beings can only choose to maximize their utility. In the framework of consumption function, individuals maximize utility that is expressed as a function of the individual's consumption stream over the span of their lifetime: $U_j = U_j\,(C_t, C_{t+1}, C_{t+2}, \ldots,$

BOX 8.4

FRANCO MODIGLIANI

Franco Modigliani (1918–2003) was born in Rome, Italy, where he completed his Bachelor degree in Economics. After emigrating to the United States, he earned his doctorate in 1944 from the New School of Social Research (now New School University). A seminal research article, which was based on his dissertation, provided the formal integration of money into the Keynesian system. In the 1950s, he suggested the life-cycle hypothesis as an alternative to the Keynesian consumption function. Modigliani also made significant contributions to the field of financial analysis and to the 'monetary/fiscal policy debate', which went on for more than 60 years. In the early 1960s he moved to Massachusetts Institute of Technology (MIT), USA, where he spent the rest of his academic career. At MIT, he designed (with Albert Ando) a large computer model of the US economy. Modigliani also used this model to attack the pro-monetarist conclusions of Milton Friedman and his disciplines.

C_L), where U_j is the utility of individual j, C_t is present consumption, C_{t+1} is next year's consumption, and so on, until the end of lifetime C_L.

The above utility function is maximized subject to the present value or worth of total resources, current and future, which will accumulate over the individual's working life or up to their retirement. These resources can be identified as the sum of the individual's present assets plus the present value of the stream of their annual disposable income until retirement. This setting implies that the individual will be able to maintain a stable pattern of consumption throughout their lifetime. In addition, income from employment will behave in a fairly predictable manner.

In Figure 8.5 the horizontal axis shows time, while the vertical axis shows income and consumption expenditures. Point 0 marks the beginning of the person's working life and T_2 marks the point of their retirement. In the early years of a person's working life and also after retirement, income is lower than consumption expenditures. For the time span T_1 to T_2 the reverse is true. As the person proceeds through life, their productivity increases and as a result their income increases up to the point of retirement T_2. From this point onwards, consumption declines, but not as much as income. The hypothesis suggests that individuals take into account the profile of the stream of income during their whole productive life and pace their consumption expenditures accordingly. At the early stages of their working life and between retirement and death they have negative savings.

Figure 8.5 The life-cycle hypothesis

These are balanced by the positive savings occurring at the time period T_1 to T_2.

The above imply that the consumption function is of the general form:

$$C_t = KV_t \tag{8.8}$$

where C_t is the current consumption by an individual, V_t is the present value of the total resources accruing to the individual over the rest of their life, and K is the factor of proportionality (meaning that consumption and V_t are connected by a constant K). In turn, the total resources available to the individual over their entire life span are the sum of the individual's net worth at the end of the previous period, plus their income during the current period from non-property sources, plus the total of the discounted values of the non-property incomes expected in the future time periods.

Assuming for simplicity that real interest rate is zero, the individual divides their resources equally over time in order to smooth out consumption: $C = (W + RY)/T$, where W is initial wealth, Y is annual income until retirement (assumed constant), R is the number of years until retirement, and T is lifetime in years. The same relation can be written as: $C = \alpha W + \beta \Upsilon$ with $\alpha = 1/T$ and $\beta = R/T$, where α is the marginal propensity to consume out of wealth, and β is the marginal propensity to consume out of income.

The life-cycle hypothesis can also account for the discrepancies in empirical consumption data discussed before. The reasoning is the following. The life-cycle consumption function divided by income can be written as $C/Y = \alpha(W/Y) + \beta$. Across households, income varies more than wealth, and this implies that high-income households should have a lower APC than low-income households. However, in the long run, aggregate wealth and income grow together and this results in a stable APC.

Apart from assuming perfectly rational consumers, the life-cycle consumption theory also assumes that individuals are indifferent to the form in which resources accrue. For example, a given increase in resources will have the same effect on consumption whether that increase takes the form of an increase in current income, expected income or net worth (Box 8.5).

One of the main criticisms in this regard is that switching of assets is not a costless transaction. Further, this theory ignores imperfections in capital markets, which impose a limit on the amount an individual can borrow (liquidity constraint). Liquidity constraints affect the ability of households

BOX 8.5

HOMO OECONOMICUS AT WORK: LIFE-CYCLE THEORY

In the consumption function given as $C_t = KV_t$, V_t can be expressed as follows:

$$V_t = \alpha x_{t-1} + Y_t + \sum_{T=1}^{n} \frac{Y^e_{t+T}}{(1 + r_t)T}$$

The above equation shows that total resources (V_t) are made up by three components. The first component is any net worth that is carried over from the previous period (αx_{t-1}). It may take the form of inherited wealth or accumulated worth. The second term is current income Y_t, and the third term is the present value of expected future income from employment Y^e over the remainder of the individual's lifetime (n).

Substituting the above into $C_t = KV_t$, we get:

$$Ct = K_t\alpha x_{t-1} + K_tY_t + K_t\sum_{T=1}^{n} \frac{Y^e_{t+T}}{(1 + r_t)T}$$

From this equation we can see that the discount rate (r) is an important element in the consumption decision. Consumers are assumed to have perfect computational abilities and to be able to estimate future income correctly (no fundamental uncertainty). The conceptual basis of this formulation is in the mainstream theory of intertemporal choice, which is based on the hypothesis of '*Homo oeconomicus*'.

to transfer resources across time periods. Finally, another line of criticism has to do with the notion of expected income, which is part of the individual's total resources. This variable is not directly observable, and its value has to be forecast, something that poses difficulties for the empirical testing of the theory (see Deaton, 1992).

Permanent income hypothesis

Apart from the life-cycle theory, the other attempt to criticize Keynes's approach to consumption was made by Milton Friedman with his permanent income hypothesis (Friedman, 1957), where permanent income is an individual's income over their lifetime. In his attempt to define a consumption function, Friedman (1957) rejects Keynes's use of current income as the determinant of consumption expenditure, based on the idea that consumers are forward-looking (meaning that future concerns affect current consumption decisions). Forward-looking consumers are a common point between Friedman's theory and the life-cycle theory. However, according to Friedman, current income is subject to random, transitory fluctuations; while according to life-cycle theory, current income changes systematically as people move through their life cycle.

Further, the permanent income hypothesis is a special case of an intertemporal optimization model of consumer behaviour, where agents maximize the sum of their expected utility subject to a lifetime budget constraint (Meghir, 2004). Consumers use their savings (or borrow) in an attempt to smooth consumption between good and bad years. These imply that current income differs from permanent income: $Y_t = Y^P + Y^T$, where Y is current income at time t, Y^P is permanent income projected at time t, and Y^T is transitory (or unexpected changes in) income. The transitory component has an expected value of zero, reflecting the notion that over time transitory gains are offset by future transitory losses, and vice versa. Thus, in the long run, observed levels of income (Y) are equal to permanent income (Y^P).

An important part of Friedman's theory was his assumption that permanent income is an average of income over the last several years. This implies that if there is a sudden rise in current income, there would be only a small increase in permanent income, contrary to Keynes's theory. Income would have to increase for several years continuously before people would expect permanent income to increase. In other words, consumers correct their previous estimates of permanent income by the amount of deviation of current income from previous period estimated permanent income (adaptive expectations).

In the same way as income, consumption (C) is divided into permanent consumption, C^P, and transitory consumption, C^T. Thus $C_t = C^P + C^T$. Like transitory income, transitory consumption is regarded as temporary. Friedman assumes that permanent consumption is a constant proportion (a) of permanent income, while permanent and transitory consumption may be interpreted as planned and unplanned consumption, respectively. Based on Friedman's assumption that Y^T is uncorrelated with C, any unforeseen increment in income does not result in unplanned consumption. Friedman justifies this premise by pointing out that even if income is other than expected, the consumer would tend to stick to their consumption plan, but adjust their asset holdings.

Given the above, Friedman's consumption function is $C = aY^P$, with $0 < a < 1$, where a is the fraction of permanent income that people consume per year. The APC will be:

$$\text{APC} = \frac{C}{Y} = \frac{aY^P}{Y} \tag{8.9}$$

Friedman's reconciliation of the empirical findings on consumption was based on the differences in consumption behaviour of different income groups. Observed short-run behaviour is explained through the value of

transitory income for different income groups. For high-income groups, transitory gains exceed transitory losses such that transitory income is on average positive over time. For low-income groups, transitory losses exceed transitory gains; while for middle-income groups the value of transitory income is equal to zero over time such that observed and permanent income take the same value. Over the long run, income variation is due mainly if not solely to variations in permanent income, which implies a stable APC.

In general, Friedman's permanent income hypothesis offered another way of explaining the conflicting results of early empirical studies on consumption. It distinguished between a short-run and a long-run consumption function. The long-run function was essentially a proportional relation, while the short-run function was a non-proportional one. The criticisms of the life-cycle theory involving the notions of liquidity constraints and of the observability and measurability of permanent income also apply to Friedman's theory.

Random walk theory of consumption and rational expectations

Contemporary mainstream theories of consumption functions are essentially extensions of the life-cycle and the permanent income theories. They are also based on Fisher's intertemporal choice model. The new element is the assumption of rational expectations: people use all available information to forecast future variables such as income. One indicative example of such theories is Robert Hall's (1978) 'random walk theory of consumption'. The rational expectations assumption means that consumption should follow a random walk: changes in consumption should be unpredictable. It also means that a change in income or wealth that was anticipated has already been factored into expected permanent income, so that it will not change consumption. Therefore, only unanticipated changes in income or wealth that change expected permanent income will change consumption. In the rational expectations framework, agents anticipate the future and therefore make all the required adjustments in the current period. The equation for future consumption is:

$$C_{t+1} = C_t + Q_{t+1} \tag{8.10}$$

In this equation, Q_{t+1} is a rational expectations error that cannot be predicted with any information known at time t. All future information is reflected in current consumption, C_t. The random walk characteristic of consumption is seen by writing $C_{t+1} - C_t = Q_{t+1}$. Consumption is a random walk, as changes over time are unforeseeable.

Clearly, the policy implications of these models are that policy changes will affect consumption only if they are not anticipated. These arguments greatly diminish the Keynesian case for government intervention, given that stabilization policies cannot be applied in any systematic way.

The heterodox perspective

Economic rationality and aggregate consumption

As mentioned above, the standard model of economic rationality is the underlying basis of all mainstream theories of consumption function. The mainstream view of economic rationality means that:

1. Consumers constantly engage in optimization under constraints. In the framework of expected utility, they are assumed to be able to assign and calculate probabilities about future decisions operating in a world of calculable probabilities.
2. Consumers have unlimited computational abilities and they are able to process unlimited information.
3. Consumers have insatiable wants, which are reducible and inseparable.
4. Consumers' preferences are made independently of those of other agents. In other words, there is no social interaction among agents as far as economic decisions are concerned.

Heterodox schools of economic thought reject this model of economic rationality and of the 'rational economic man' (*Homo oeconomicus*). First of all, heterodox economists believe that, in most of consumers' decisions, optimizing is impossible either because of a lack of information or because of an overload of information and deficient computational capabilities. Instead, they follow the model of procedural or bounded rationality initiated by Nobel Prize winner Herbert Simon (Simon, 1979; Box 8.6).

In the same vein, and following Keynes's approach to uncertainty, heterodox theorists argue that the existence of fundamental uncertainty further undermines the optimization process and especially the expected utility theory.

The financial crisis of 2007–08 offered many examples of 'fundamental uncertainty' illustrated by unlikely events ('black swans'; Box 8.7) incompatible with normal probability distributions assumed by the expected utility theory.

Consequently, and contrary to the assumptions of mainstream consumption theories, future incomes (and also costs and revenues) are almost impossible

BOX 8.6

HERBERT SIMON

Herbert Simon (1916–2001) was an American economist who won the Nobel Prize in Economics in 1978 for his studies of administrative behaviour and decision-making within large organizations. He earned a PhD from Harvard University in political science in 1943. Apart from economics, his research interests covered a wide range of subjects such as cognitive science, computer science, public administration, management and political science. Among economists, he is best known for his revolutionary work in microeconomics and especially for his critique of the mainstream model of '*Homo oeconomicus*'. Simon argued that the conception of economic man as a 'lightning calculator' of costs and benefits is unrealistic. Instead, he suggested that the costs of obtaining information about alternative outcomes, and the presence of uncertainty about the future, lead to bounded rationality. This means that agents are unable to engage in constant choice optimization but engage in satisficing: they try to achieve certain aspiration levels or targets, and they gradually adjust these levels up or down depending on how close they are to the original target.

Note: The phrase 'lightning calculator' originates from Veblen's (1898, p. 389) critique of *Homo economicus*: 'The hedonistic conception of man is that of a lightning calculator of pleasures and pains, who oscillates like a homogeneous globule of desire of happiness under the impulse of stimuli that shift him about the area but leave him intact.'

BOX 8.7

BLACK SWANS

The black swan theory is used to describe events that deviate beyond what is normally expected of a situation. Black swan events are typically random and unexpected, and are extremely difficult to predict. The event generally has a major effect or is shocking in nature, and comes as a complete surprise. The theory was pioneered by the finance professional Nassim Nicholas Taleb, and it is a combination of mathematical and philosophical reasoning to explain and describe the randomness of uncertainty. The financial crash of the US housing market during the 2008 crisis is one of the most recent and well-known black swan events.

to predict. In this framework, and following Keynes's analysis of animal spirits, consumers' decisions depend on 'spontaneous optimism' or rules of thumb rather than calculations of (expected) costs and benefits (see Lavoie, 1994; Davidson, 2009; King, 2015).

In contrast to mainstream theory, therefore, heterodox economists distinguish between wants and needs, clearly implying the necessity aspect of human needs. Moreover, the principle of separability of needs says that needs can be distinguished from each other and that there exists a hierarchy of needs, in which basic needs are satisfied first, and non-basic needs come into the picture only once basic needs are satisfied (Lavoie, 1994).

Finally, heterodox economists argue that economic decisions are interdependent. This suggests that consumers care about their relative position and compare their income, wages or wealth to other people's in their social reference group. This idea can also be found in Keynes's notion of relative real wage: 'workers resist a cut in money wages in order to maintain their relative position in the wage structure and not so much to avoid a cut in their absolute income' (Keynes, 1936, p. 14).

There is also another important point that has to do with the necessity of microeconomic foundations (see the Conclusion of this volume). All post-Keynesian economists agree that it is impossible to base macroeconomic theory on representative agents with rational expectations. The insistence on providing rigorous neoclassical microfoundations for macroeconomic theory amounts to a denial of the fallacy of composition, or gives rise to paradoxes (see Chapter 1), which Keynes regarded as the methodological precondition for having a separate macroeconomics (King, 2015).

In view of the above, the rational-agent basis of mainstream consumption theories is rejected. Instead, a theory of aggregate consumption behaviour must reflect that consumers do not engage in optimizing behaviour. It should also take into account that consumption decisions are psychology-driven and are made in a social context in which consumers care about their social position and social status. One indicative example of such theory was suggested by James Duesenberry, to which I now turn (see his portrait at the end of this chapter).

The relative income hypothesis

In 1949, Harvard professor James Duesenberry made the very first attempt at providing a theoretical justification for the discrepancy between Kuznets's short-run and the long-run empirical findings on consumption. Duesenberry suggested the relative income hypothesis as the main theory underlying the consumption function. The relative income hypothesis introduces psychological and sociological factors such as social interdependencies and habit formation to the study of consumer behaviour. The emphasis on the

social dimension of consumption was not a new idea. The idea that people compare their income, consumption or wealth with other people's income, consumption or wealth can be found in the work of many major economists such as Adam Smith, Karl Marx, John Stuart Mill, Thorstein Veblen and Arthur Cecil Pigou (see Drakopoulos, 2016). Duesenberry, however, was the first to apply the concept of social comparisons to the study of consumption in a systematic manner. In fact, Duesenberry's work can be viewed as a continuation of Veblen's ideas, given that there are many common points concerning income and consumption comparisons, and also concerning the role of the demonstration effect (see below). Further, the common point of Duesenberry's and Keynes's approaches is the idea of social comparisons or relative standing: Duesenberry put emphasis on relative consumption, while Keynes emphasized relative wage. Although Keynes recognized the importance of social influences on consumption decisions, he did not develop them further in his *General Theory*, arguing that they were stable, at least in the short run (see Keynes, 1936; Mason, 2000). On the contrary, in his consumption theory, Duesenberry provided ample analytical insights regarding social comparisons.

In his 1949 book entitled *Income, Saving and the Theory of Consumer Behavior*, Duesenberry starts by arguing that two 'fundamental assumptions' of demand theory are 'invalid'. These assumptions are '(1) that every individual's consumption behaviour is independent of that of every other individual, and (2) that consumption relations are reversible in time' (Duesenberry, 1949, p. 1). Let us first look at the implications of assumption (1) and then those of assumption (2).

Duesenberry (1949, p. 3) starts by arguing that the assumption of independent preferences has 'no empirical basis' and then states that 'there are strong psychological and sociological reasons for supposing that preferences are in fact interdependent'. In the context of consumption theory, this meant that consumers will be influenced by the behaviour of other consumers: 'any particular consumer will be influenced by consumption of people with whom he has social contacts' (ibid., p. 48). This idea (labelled as the 'demonstration effect') questioned the established view that absolute levels of income only determine patterns of consumer demand (ibid., p. 27). Consequently, he maintained that a household's consumption would depend not just on its own current level of income, but also on its income relative to those in the subgroup of the population with which it identifies itself. The household will attempt to match its consumption behaviour and thus its consumption expenditures with those of other members of its group. It follows that households with lower income within the group will consume a larger share of their

income to 'keep up with the Joneses', while households with high incomes relative to the group will save more and consume less. As Duesenberry (1949, pp. 27–8) writes: 'We can maintain then that the frequency and strength of impulses to increase expenditure depends on frequency of contact with goods superior to those habitually consumed. This effect need not depend at all on considerations of emulation or "conspicuous consumption"'.

The frequency and strength of impulses to increase expenditure depend on frequency of contact with goods superior to those habitually consumed. Duesenberry (1949, p. 27) appeals to self-observation in order to support the presence of the demonstration effect:

> The best way to demonstrate that consumption expenditures can be forced up by contact with superior consumption goods is to ask the reader to consult his own experience. What kind of reaction is produced by looking at a friend's new car or looking at houses or apartments better than one's own? The response is likely to be a feeling of dissatisfaction with one's own house or car. If this feeling is produced often enough it will lead to action, which eliminates it, that is, to increase expenditure.

When consumers frequently come into contact with superior goods, they are constantly reminded of their low social status. The result will be 'an increase in expenditure at the expense of saving' (ibid., p. 27). Thus, households with lower income within the group will consume a larger share of their income to 'keep up with the Joneses', while households with high incomes relative to the group will save more and consume less. The demonstration effect can be understood better with reference to the two short-run consumption function flatter lines labelled C_{SR1} and C_{SR2} in Figure 8.4. Line C_{SR1} might represent the cross-section consumption function of a low-income group, and line C_{SR2} might represent the high-income group. As incomes of both groups rise over time, both lines would tend to slide up the steeper 'long-run consumption function', with the average household in each group tending to spend a constant share of its income over time. Figure 8.6 explains Duesenberry's approach to reconcile the two sets of data.

In Figure 8.6, C_m shows family consumption expenditures and Y_m shows family real disposable income. Consumption schedules I, II and III indicate the relationship between family income and family consumption at times 1, 2 and 3, respectively. Consumption schedule I depicts this relationship at time $t = 1$. The shape of this family consumption shows that within a distribution of income, the APC falls as we move from low-income to high-income families. This feature is consistent with both the Keynesian absolute income

Figure 8.6 Duesenberry's long-run and short-run consumption functions

hypothesis and Duesenberry's relative income hypothesis. If we assume that all families receive proportional increases in income in period $t = 2$, the distribution of income remains the same. In terms of the absolute income hypothesis, this would mean a movement along consumption schedule I. In terms of relative income, this would mean that the consumption schedule will shift from I to II. More specifically, assume that in $t = 1$ mean family income is Y_{m1} and the corresponding level of consumption is C_{m1}. After the proportional increase in all incomes, the family will maintain its position in the income distribution by earning Y_{m2} in $t = 2$, thus its APC will remain unchanged resulting in a level of consumption at C_{m2}. The extension of this principle to all families gives consumption schedule II. If the same process is repeated in $t = 3$, we would derive consumption schedule III. In sum, a general proportional rise in the levels of income would cause successive cross-section studies to indicate shifts in family consumption schedule. As the consumption schedule shifts over long periods of time and with unchanged income distribution, the long-run time series will indicate a proportional relationship between aggregate consumption and aggregate income: $C_t = kY_t$, with $APC = \frac{Ct}{Yt} = k$. In other words, the APC is constant in the long run, in accordance with Kuznets's findings (see also Venieris and Sebolt, 1977).

The second basic component of the relative income hypothesis is that 'consumption relations are reversible in time' (Duesenberry, 1949, p. 1). In his own words: 'Over time, the relation between aggregate consumption and aggregate income is not completely reversible. As income increases secularly,

consumption will grow proportionally; but over the cycle, as income falls from its peak, consumers will attempt to maintain consumption standards set previously' (Duesenberry, 1949, p. 7).

The main difference from the demonstration effect is that instead of comparing their income to those of other households, each household is assumed to consider its current income relative to its own past income levels. A household that has in the past achieved income levels higher than its present levels would attempt to maintain the high consumption levels that it achieved earlier. Thus, when incomes fall, consumption would not fall in proportion. The result of this behaviour for aggregate consumption is also called a 'ratchet effect'.

In Duesenberry's work, therefore, aggregate consumption depends not only on current disposable income but also on the ratio of current income to previous peak level income. The consumption function in terms of the average propensity to consume (C_t/Y_t) is written as:

$$\frac{C_t}{Y_t} = a - b\left(\frac{Y_t}{Y_o}\right) \tag{8.11}$$

where C_t is current consumption, Y_t is current disposable income, and Y_0 is previous peak income.

By referring to Figure 8.6, when incomes rise, consumption increases along the steeper long-run consumption function. However, when a recession hits and incomes decline, households reduce consumption less than proportionally and fall back along the flatter short-run consumption function. During the recovery, they move up along the flat line until they reach their highest attained level of consumption. After recovery, when incomes grow again, they proceed up the long-run line again until the next recession, when they fall back along a flatter line. Thus, consumption ratchets upward, staying relatively near its highest past value when income declines. The implication is that in the long run APC will be constant, but as the economy moves through the business cycle, the ratio of current to previous peak income will vary and thus APC will also follow the cyclical fluctuations (Box 8.8).

Duesenberry's theoretical approach was able to reconcile the discrepancy between the empirical cross-section studies and the long-run findings. The main theoretical implication, namely that the APC will be greater than the MPC, lies solidly within the Keynesian tradition (Kosobud, 1998, p. 135). Further, Duesenberry's theory suggests that fiscal changes may have an

BOX 8.8

APC, MPC AND THE RELATIVE INCOME HYPOTHESIS

Given that consumption is also related to previous peak income, the consumption function is given as:

$$Ct = \alpha Yt - b\left(\frac{Yt^2}{Yo}\right)$$

Thus:

$$APC = \frac{Ct}{Yt} = \alpha - b\left(\frac{Yt}{Yo}\right)$$

If income grows over the long run at an average rate of g percent per year, the long-run ratio of current income to previous peak income would be:

$$\frac{Yt}{Yo} = (1 + g)$$

Then the APC will be:

$$\frac{Ct}{Yo} = \alpha - b(1 + g) = k$$

where k is a constant. Further, the MPC is:

$$MPC = \frac{\partial C}{\partial Y} = \alpha - 2b\left(\frac{Yt}{Yo}\right) = \alpha - 2b(1 + g)$$

A comparison shows that MPC < APC, which is in agreement with Keynes's views.

asymmetrical effect. Tax reductions may well stimulate consumption spending. However, tax increases may not have a significant effect in curbing demand in the short run, as consumers strive to maintain consumption levels.

Policy implications and concluding remarks

In comparison to Keynes and Duesenberry, mainstream consumption functions have very different consequences for economic policy prescriptions. For example, in Friedman's framework of permanent and transitory components, a much larger part of current consumption is considered as autonomous, and a much smaller part as dependent on current income. Since the marginal propensity to consume from transitory income is zero, increases in income arising from increases in government spending and/or falling taxes will have negligible effects on the economy. The fiscal multipliers (assuming the change is viewed as temporary) will be small or even zero. Given the relationship between current consumption and the magnitude of the fiscal multipliers, Friedman's theory implied smaller fiscal multipliers and thus a largely ineffective fiscal policy. It also implied an inherently more stable economic system (Bunting, 1989).

Similar observations hold for the life-cycle hypothesis. Although the theory takes into consideration the role of current income, it places greater emphasis than the Keynesian approaches on the role of expected income and wealth

on consumption decisions. For instance, changes in current income arising from fiscal policy will have a strong effect on current consumption only if they affect expected lifetime income. Thus, the points regarding the role of the associated multipliers and fiscal policy made in reference to the permanent income hypothesis also apply here.

Contemporary mainstream macroeconomic models employ consumption functions that combine rational expectations with forward-looking agents. In these models, the role of economic policy is very limited, given that policy changes will affect consumption only if they are not anticipated. However, the validity of these models (especially after the Great Recession of 2008) has been questioned even by leading mainstream theorists in terms of 'unrealistic microfoundations for the behavior of households embodied in the "rational expectations permanent income" model of consumption' (Muellbauer, 2016, p. 2).

By greatly diminishing the role of current income in consumption functions, mainstream consumption theories equally diminish the role of fiscal policy in preventing economic downturns and thus employment fluctuations. In heterodox theories where current or relative income is a major determinant of consumption, changes in income will bring changes in consumption. These changes will be large and occur within a short time span, and this means that fiscal policy can be used as a major instrument in order to curb unemployment and economic recessions.

REFERENCES

Ackley, G. (1960), *Macroeconomic Theory*, New York: Macmillan.

Bunting, D. (1989), 'The consumption function "paradox"', *Journal of Post Keynesian Economics*, **11** (3), 347–59.

Davidson, P. (2009), *The Keynes Solution: The Path to Global Economic Prosperity*, New York: Palgrave Macmillan.

Deaton, A. (1992), *Macroeconomics and Consumption*, Oxford: Oxford University Press.

D'Orlando, F. and E. Sanfilippo (2010), 'Behavioral foundations for the Keynesian consumption function', *Journal of Economic Psychology*, **31** (6), 1035–46.

Dow, S.C. (1985), *Macroeconomic Thought: A Methodological Approach*, London: Basil Blackwell.

Dow, S.C. and J. Hillard (eds) (1995), *Keynes, Knowledge and Uncertainty*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Drakopoulos, S.A. (1990), 'Two levels of hedonistic influence on microeconomic theory', *Scottish Journal of Political Economy*, **37** (4), 360–78.

Drakopoulos, S.A. (1992), 'Keynes's economic thought and the theory of consumer behaviour', *Scottish Journal of Political Economy*, **39** (3), 318–36.

Drakopoulos, S.A. (2016), *Comparisons in Economic Thought: Economic Interdependency Reconsidered*, London, UK and New York, USA: Routledge.

Duesenberry, J.S. (1949), *Income, Saving and the Theory of Consumer Behavior*, Cambridge, MA: Harvard University Press.

Fisher, I. (1892/1961), *Mathematical Investigations in the Theory of Value and Prices*, New York: A.M. Kelley.

Fisher, I. (1930), *The Theory of Interest*, New York: Macmillan.

Friedman, M. (1957), *The Consumption Function*, Princeton, NJ: Princeton University Press.

Hall, R.E. (1978), 'Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence', *Journal of Political Economy*, **86** (6), 971–87.

Keynes, J.M. (1936), *The General Theory of Employment, Interest, and Money*, London: Macmillan.

Keynes, J.M. (1978a), *The Collected Writings of John Maynard Keynes, Vol. VIII: A Treatise on Probability*, London: Macmillan.

Keynes, J.M. (1978b), *The Collected Writings of John Maynard Keynes, Vol. X: Essays in Biography*, London: Macmillan.

King, J.E. (2015), *Advanced Introduction to Post Keynesian Economics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Koçkesen, L. (2008), 'The relative income hypothesis', in W.A. Darity (ed.), *International Encyclopedia of the Social Sciences, 2nd edition*, Detroit: Macmillan, pp. 153–4.

Kosobud, R. (1998), 'Relative income hypothesis', in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave: A Dictionary of Economics*, Vol. 4, London: Macmillan, pp. 134–6.

Kuznets, S. (1946), *National Product Since 1869* (assisted by L. Epstein and E. Zenks), New York: National Bureau of Economic Research.

Lavoie, M. (1994), 'A post Keynesian theory of consumer choice', *Journal of Post Keynesian Economics*, **16** (4), 539–62.

Lawson, T. (1988), 'Probability and uncertainty in economic analysis', *Journal of Post Keynesian Economics*, **11** (1), 38–65.

Mason, R. (2000), 'The social significance of consumption: James Duesenberry's contribution to consumer theory', *Journal of Economic Issues*, **34** (3), 553–72.

Meghir, C. (2004), 'A retrospective on Friedman's theory of permanent income', *Economic Journal*, **114** (496), F293–F306.

Modigliani, F. and R. Brumberg (1954), 'Utility analysis and the consumption function: an interpretation of cross-section data', in K.K. Kurihara (ed.), *Post-Keynesian Economics*, New Brunswick, NJ: Rutgers University Press, pp. 388–436.

Muellbauer, J. (2016), 'Macroeconomics and consumption', Centre for Economic Policy Research Discussion Paper, No. 11588, and Oxford University Department of Economics Working Paper, No. 811.

Savage, L. (1962), *The Foundations of Statistical Inference*, London: Methuen & Co.

Simon, H. (1979), 'Rational decision making in business organizations', *American Economic Review*, **69** (4), 493–513.

Veblen, T. (1898), 'Why is economics not an evolutionary science?', *Quarterly Journal of Economics*, **12** (4), 373-97.

Venieris, Y. and F. Sebold (1977), *Macroeconomics: Models and Policy*, Santa Barbara: John Wiley & Sons.

A PORTRAIT OF JAMES STEMBLE DUESENBERRY (1918–2009)

James Stemble Duesenberry was born in Princeton, West Virginia, United States, on 18 July 1918. He graduated from Bennett High School in Buffalo, New York, and attended the University of Michigan, where he completed his undergraduate degree in 1939. Duesenberry received his Masters degree in 1941 and his doctorate in 1948 from the same university. He initially worked as a teaching assistant at the University of Michigan and became an instructor at the Massachusetts Institute of Technology (MIT) in 1946. In the same year, he also began his Harvard career as a teaching fellow. He was also a Fulbright fellow at the University of Cambridge, UK in 1954–55, before being appointed as a full Professor of the Harvard Faculty in 1955. He was a Ford Research Professor in 1958–59, and he continued teaching at Harvard until 1989. Until his death, he was Emeritus Professor in the Department of Economics at Harvard University. During his career, Duesenberry served as a member of the President's Council of Economic Advisors under President Johnson from 1966 to 1968.

Duesenberry's most famous contribution was his book entitled *Income, Saving and the Theory of Consumer Behavior*, which was based on his doctoral dissertation and was published one year after the completion of his degree (1949). The relative income hypothesis is the central thesis of this book. The hypothesis was able to account for both the cross-sectional and time series evidence and was an extension and improvement of Keynes's consumption theory. According to Duesenberry, the utility that individuals derive from consumption depends on what and how much the individual consumes relative to what and how much others consume. The underlying idea here is that people are concerned about their social status relative to others. It is clear that the relative income hypothesis emphasizes the social nature of consumption patterns.

The social nature of decisions was also part of Keynes's economic approach. According to Keynes, the failure of money wages to fall in times of massive unemployment can be understood as an effort of the workers to protect their relative real wage. It follows that the common ground of Duesenberry's and Keynes's approaches was the idea that individuals do not operate in social isolation, but compare their economic situation and their relative social standing to other similar individuals. In Duesenberry's work, social comparisons occur mainly through the concept of relative consumption, while in Keynes's work they occur through the concept of relative wage. As a conceptual continuation of Keynes's ideas, Duesenberry provided a detailed analysis of the effects of social comparisons in consumption theory.

Duesenberry's starting point was that consumers are influenced by the consumption pattern of people with whom they have social contacts. Consequently, consumption patterns are influenced by both the frequency of contact of a household with 'superior goods' (that is, the 'demonstration effect') and a social 'drive towards higher living standards' (the 'social significance of consumption'). This approach questioned the established view that absolute levels of income only determine patterns of consumer demand. It also indicated

➡

that the aggregate saving ratio was independent of the absolute level of income. Duesenberry proceeded further to analyse the basis of such behaviour. In his view, the frequency and strength of impulses to increase expenditure depends on frequency of contact with goods superior to those habitually consumed. In order to provide supplementary foundations to the notion of relative consumption, he suggested self-observation to experience the relative deprivation of the demonstration effect. He referred to the feelings of dissatisfaction that someone experiences when comparing their own car or apartment to those of their friends.

The relative income hypothesis has been used to explain inconsistent empirical evidence between the cross-sectional and time series data. Cross-sectional data show that the share of income saved rises with income, while time series data show that as the economy grows and incomes rise, the share of national income saved stays more or less constant. Low-income families feel constant pressure to spend more, and thus they save less. High-income families feel less pressure to consume, and thus they save more. However, this pattern does not hold in the long run, given that the share of national income saved does not increase. The inconsistency lies in the fact that if high-income families save a higher fraction of income than low-income families, the saving rate should rise as national income becomes higher.

Duesenberry's relative income hypothesis resolves the inconsistency. Concern with relative income means that aggregate saving rate is independent of aggregate income. For example, if everybody's income has increased, the demonstration effect will be just as powerful as it was before, and the saving rate will be the same. This is consistent with the time series evidence. The relative income hypothesis also means that the propensity to save of an individual is an increasing function of their percentile position in the income distribution. This is consistent with the cross-sectional evidence.

Another important economic implication of the relative income hypothesis is that consumption creates negative externalities in the society, much in the same way that pollution does. Individuals who care about their relative income will want to consume and work more in order to increase their social status. It is very likely that in order to achieve this, they will tend to work too much relative to the socially optimal level. As in the case of pollution, an appropriate income taxation policy could improve social welfare.

The relative income hypothesis was a very significant contribution because of its realistic approach and its ability to explain the available empirical evidence. Therefore, it was not surprising that soon after its appearance, Duesenberry's work was acknowledged as a very important contribution by many leading economists of the period. Kenneth Arrow, George L.S. Schackle and Artur C. Pigou expressed positive views about the relative consumption hypothesis. Similarly, the position of leading consumption theory specialists such as Franco Modigliani and Richard Brumberg, towards Duesenberry's analysis, was initially sympathetic.

However, a few years later, Modigliani and Brumberg claimed that their new interpretation of consumption theory was sounder and much simpler than Duesenberry's. Modigliani argued that it

contained unnecessary social and psychological elements. During the same period, Milton Friedman also attacked Duesenberry's formulations by declaring that permanent income rather than relative income was the basis of consumer behaviour. Although Friedman acknowledged some merit in Duesenberry's concept of relative income, he argued that it was basically only a biased index of relative permanent income status. Furthermore, Friedman believed that his approach was superior, since it owed nothing to sociology or to psychology, in contrast to Duesenberry's, which was full of subjective elements.

Much in the same way as the previous approaches involving the social context of economic decisions, Duesenberry's ideas never gained popularity among mainstream theorists. One reason has to do with the policy implications of the relative income hypothesis. Similarly to Keynes's theory, Duesenberry's approach provides justification for the use of fiscal policy to curb unemployment and economic recessions. This view is at odds with the mainstream position of the very limited effectiveness of fiscal policy in a free market economy. The other reason might be that Duesenberry's analysis poses some fundamental questions regarding the cultural influences on economic decisions and the endogeneity of preferences, issues which were not (and still are not) popular for orthodox theorists. Thus, it seems that allegedly simpler and non-psychological explanations of consumption patterns offered by Modigliani, Brumberg and Friedman were preferred to Duesenberry's more sophisticated approach. Although even today some well-known theorists such as Robert H. Frank argue that it outperforms the mainstream consumption theories, Duesenberry's consumption theory was effectively rejected because of its non-mainstream methodological foundations.

EXAM QUESTIONS

True or false questions

1. According to the permanent income hypothesis, consumers correct their previous estimates of permanent income by the amount of deviation of current income from previous period estimated permanent income.
2. According to the random walk theory of consumption, anticipated changes in income or wealth will change consumption.
3. James Duesenberry maintained that during an economic downturn consumers will attempt to maintain consumption standards set in previous good years.
4. The notion of Keynes's fundamental uncertainty does not undermine the rational agent, forward-looking-based theories of consumption.
5. In the Keynesian system, the marginal propensity to consume (MPC) determines the magnitude of government and tax multipliers.
6. In the framework of intertemporal choice theory of consumption, the optimal point for the consumer is where the slope of the budget line just touches the lowest possible indifference curve.
7. Both the absolute income hypothesis and the relative income hypothesis imply that the average propensity to consume will be lower than the marginal propensity to consume.

8. In the intertemporal choice theory of consumption, the level of interest rate plays an important role in consumption decisions.
9. The model of procedural or bounded rationality implies that consumers' decisions exhibit optimizing behaviour.
10. The life-cycle consumption theory contains the notion of the marginal propensity to consume out of wealth.

Multiple choice questions

1. According to the intertemporal choice theory of consumption:
   a) consumers are rational, but backward-looking agents;
   b) consumers choose consumption levels for the present and future so as to maximize lifetime satisfaction;
   c) consumers consider only current income for their consumption choice;
   d) consumers consider only relevant income for their consumption choice.
2. According to Keynes:
   a) future expected income determines current consumption;
   b) the level of wealth determines current consumption;
   c) current disposable income determines current consumption;
   d) permanent income determines current consumption.
3. The Keynesian consumption function implies that:
   a) households with higher income will consume less, will save more, and that APC will be falling as income increases;
   b) households with higher income will consume more, will save more, and that APC will be increasing as income increases;
   c) households with higher income will consume less, will save less, and that APC will be falling as income increases;
   d) households with higher income will consume more, will save more, and that APC will be falling as income increases.
4. In the context of the life-cycle consumption function:
   a) consumers will be able to maintain a stable pattern of consumption throughout their lifetime;
   b) consumers will experience varying levels of consumption throughout their lifetime;
   c) consumers are only interested in consumption and not savings;
   d) consumers will always have higher income than their consumption expenditures.
5. Duesenberry's demonstration effect means that:
   a) a household's consumption would only depend on its income relative to those in the subgroup of the population with which it identifies itself;
   b) a household's consumption would depend on its own current level of income and on its income relative to those in the subgroup of the population with which it identifies itself;
   c) a household's consumption would depend only on its own current level of income;
   d) a household's consumption would depend on its permanent level of income.
6. If MPC = 0.6, the magnitude of private investment and government spending multiplier is:
   a) equal to 10;
   b) equal to 8;
   c) equal to 2.5;
   d) equal to 6.

7. When consumption is a function of current and relative income, fiscal policy is more likely to be:
   a) ineffective in reducing unemployment;
   b) ineffective in smoothing economic downturns;
   c) irrelevant for economic policy;
   d) effective in reducing unemployment.
8. According to mainstream consumption theories, an increase in government spending will most likely:
   a) have a positive effect on aggregate income;
   b) affect savings only;
   c) have no effect on aggregate income;
   d) none of the above.
9. In the framework of the expected utility approach, probabilities are:
   a) based on frequency and are numerically measurable;
   b) infrequent;
   c) not well defined;
   d) always numerically indeterminate.
10. In which of the following notions are economic decisions are assumed to be interdependent?
   a) Relative consumption.
   b) Relative wage.
   c) Keeping up with the Joneses.
   d) All of the above.

9

Theories of investment

Thomas Dallery

OVERVIEW

This chapter:

- presents the mainstream view of investment, which rests upon the role of price factors and the principle of substitution between production factors;
- presents Keynes's view of investment while emphasizing the role of uncertainty, expectations and effective demand;
- offers a synthesis of heterodox theories of investment where profitability, demand and financing constraints contribute in shaping firms' decisions;
- gives a hint of the empirical evaluation of the factors determining investment, especially as far as financialization is concerned.

KEYWORDS

- **Accelerator models:** These models assume that investment is determined by the expected demand for firms' products. The variations in demand may imply amplified variations in investment.
- **Capacity effect:** This effect indicates that investment allows for an increase in firms' productive capacity. It is the supply side of investment.
- **Demand effect:** This effect refers to the fact that investment increases demand for firms producing capital goods. This first increase in demand is amplified owing to the multiplier principle.
- **Multiplier principle:** This principle states that an increase in investment leads to a bigger increase in output.
- **Oscillator model:** The interplay between the accelerator and multiplier principles may produce cycles in investment and activity.

Why are these topics important?

The act of investment necessarily implies the passage of time, since it deals with speculation in its essential form: an action undertaken today, the rewards of which only arise sometime in the future. All types of investment (Box 9.1) suffer from this discrepancy in time.

Physical investment designates the purchase of new equipment by a firm: they are goods that are used in the production of other goods. As for what motivates a firm to invest, we can count three possibilities: (1) new equipment may be acquired to replace old equipment that is undermined by obsolescence and/or wear and tear (this does not really increase the firm's productive capacities); or (2) it also may be bought to add to the existing stock of equipment, thereby increasing a firm's productive capacities; this is done with the objective of gaining control over markets and gaining market shares; and finally, (3) it may be bought specifically to reduce firms' costs of production, while eventually replacing workers with machines.

Yet, these three different types of physical investment rest on a similar premise: the fact that firms need to act now so as to face the future and eventual flaws

BOX 9.1

THE DIFFERENT TYPES OF INVESTMENT

Investment is a polysemous term. The same word may be used to describe different situations such as a firm buying a new machine (productive investment), a student deciding to go to college (investment in human capital), the acquisition of a stock (financial investment), or the improvement of the brand image through an advertising campaign (investment in intangible capital). Obviously, these different types of investment do not impact upon the economy in the same way. For example, when a firm undertakes a productive investment, it adds new equipment to the productive capacity of the economy. A firm that proceeds to a financial investment while buying a competitor on the stock market does not contribute to an increase in the productive capacity of the economy. Regarding capital accumulation, the distinction between gross investment and net investment is also important. While net investment alludes only to the increase in capital stock, gross investment includes the total spending in equipment goods, independently of whether they serve to replace old equipment or allow for an increase in firms' productive capacities. For developed countries, gross investment consists mainly in replacing old equipment: between 1970 and 2019, net investment represented less than a third of gross investment for the private economy in the United States and in France.

BOX 9.2

THE KALDOR–VERDOORN LAW

The Kaldor–Verdoorn Law is a centrepiece of studies on structural change. It stems from the original work of the Dutch economist Verdoorn (1949), and was later extended by Kaldor (1966) with a focus on exports. While neoclassical growth models assume that productivity growth comes from supply factors, the Kaldor–Verdoorn Law posits that productivity growth is fostered by demand factors through a cumulative causation process: an increase in expected demand leads to an increase in investment, allowing for an increase in productivity that can be transformed in price cuts, which finally contributes to a new cycle of increased demand on the product market.

in their existing capital stock. Moreover, investment is undertaken by firms to face anticipated future surge in demand on the market for produced goods and services, and/or to deal with the future pressure from competitors to become more cost-efficient.

Investment also serves the purpose of giving firms access to new technologies, thereby participating in the spreading of innovations. Capital accumulation – that is, the increase in the stock of investment equipment – may be the essential factor in cumulating technical progress, thereby inducing development (Box 9.2).

Investment is at the core of economic dynamics. It is through investment that firms prepare the future of production; it is through investment that firms implement innovations; and it is through investment that firms compete with one another. Yet, investment is precisely what makes the role of entrepreneurs so difficult. Investment is an expenditure in the hope of earning an uncertain profit in the future. Consequently, since we are dealing with the passage of time and an unknown future, radical uncertainty is at the heart of heterodox investment theories, thus following the path laid out by Keynes (1936).

Investment is also an ambivalent variable, since it affects both demand and supply. On the one hand, investment adds new productive capacities, and thus contributes to enhance and/or improve supply on the market for produced goods and services. On the other hand, investment is a source of demand for those firms producing investment goods. The interplay between these two components of investment is likely to induce instability in our economies, and it is all the more important to understand better how

investment decisions are made, and how they contribute to shape the economic system as a whole.

The traditional mainstream view

In order to deal with investment theories in neoclassical economics, one is first obliged to go back to the question of production. The common ground for neoclassical economists in this regard is the use of production functions. These may exhibit different mathematical formulations, but they all amount to the fact that a firm has to combine different production factors (most of the time only capital and labour) so as to produce goods and services. To put it bluntly, production requires both machines and workers. More precisely, since neoclassical economists assume factor substitutability, one can choose to replace workers by machines (or the other way round). For example, if wages become too high, neoclassical theory assumes that firms can simply replace workers with machines, hence why it assumes 'factor substitutability'.

Obviously, this assumption of factor substitutability opens many different ways to produce the same quantity of goods and services: some of them will use many workers and few machines (labour-intensive), some others will use few workers and many machines (capital-intensive). Firms are assumed to maximize profits, so that the choice of technique (labour-intensive or capital-intensive) adopted by the firm will depend on the costs of the different production factors. If the relative price of capital compared to labour increases, firms will turn to more labour-intensive production techniques. Respectively, if wages are going up relative to capital prices, firms will try to substitute machines for workers. To complete this rough summary of the neoclassical framework, it is just necessary to add that the price of capital is the interest rate. Consequently, when interest rates increase, there is a drop in investment. This is what one may find in most macroeconomic models of neoclassical economics.

This global picture of investment theories in neoclassical economics deserves to be developed on several points. The canonical model of investment in neoclassical economics is generally agreed to have been developed by American economist Dale Jorgenson in his 1963 paper. A whole literature has extended this model in various directions to answer the initial criticisms of the paper. Nevertheless, Jorgenson (1963) is still a good starting point to present the neoclassical theory of investment. A first remark here is to note that in Jorgenson (1963), investment is only a side-variable. Indeed, what Jorgenson (1963) tries to determine is the optimal value of the capital stock for a representative firm; as a result, investment is only the annual flow that

the firm has to spend in order for it to make its actual capital stock converge to its optimal level.

In the canonical model, firms do not undergo costs to adjust their capital stock (beyond the price of investment goods). This assumption has been criticized, because, for example, the installation of new equipment may divert workers from the activity of production, or it may imply the need for retraining workers. It has also been put forward that investment may be irreversible, so that the adjustment between the actual and optimal capital stock may be costly. Indeed, some productive equipment is specific to a firm or to an industry, so that it cannot be sold easily. The absence of adjustment costs has also been questioned on the basis that investment takes time. As a result, it is possible that between the moment when the investment decision is made and the moment when new equipment is installed, either the technology has evolved, rendering new equipment partly obsolete (accelerated depreciation), or the demand on the product market has evolved, making new equipment unnecessary.

The neoclassical literature has tried to address these different challenges in subsequent works, and nowadays the discussion revolves around an improved version of Jorgenson (1963) with the inclusion of adjustment costs. Consequently, firms do not try to fill instantaneously the gap between the actual capital stock and the optimal one, and the speed of adjustment now becomes of great interest for this literature (Box 9.3), which can be divided between those studies assuming convex adjustment costs (Hayashi, 1982) and those supposing non-convex adjustment costs (Dixit and Pindyck,

BOX 9.3

CONVEX AND NON-CONVEX ADJUSTMENT COSTS

The literature on convex adjustment costs assumes that the more a firm invests, the more it has to face costs. More precisely, these costs are increasing with investment, but they are growing at an increasing rate: it is more costly to build two small new plants than one big plant. Non-convex adjustment costs may refer to fixed costs: every time a firm desires to modify its capital stock, it may face adjustment costs that do not depend on the scale of investment spending. While convex adjustment costs theories predict smooth and continuous changes in investment to help in filling the gap between effective and optimal capital stock, non-convex adjustment costs theories are more consistent with waves of investment where firms suddenly spend a lot in capital goods after periods of no investment at all.

1994). In the end, the essential result remains the same: investment is negatively correlated to interest rates.

Let us return to the relationship between interest rates and investment. More precisely, to better understand why an increase in interest rates leads to a drop in investment, it is necessary to explore the firms' rationality. According to an old paper by Irving Fisher (1930), the underlying assumption is that firms consider a list of investment projects whose expected profitability is known (or thought to be known). It is therefore possible to rank these projects by profitability. The comparison between the expected return for investment projects and the current interest rate becomes straightforward: on the one hand, if one chooses to think in terms of opportunity cost, it is logical for a firm with a given amount of funds to invest only in projects that deliver a return that is higher than the interest rate it could gain by lending these funds; on the other hand, if one chooses to think in terms of the financial burden of indebtedness, it is also perfectly consistent for a firm with no cash to invest only in projects whose return is sufficient to at least compensate the interest rate it has to pay in order for it to borrow the funds necessary to finance these investment projects. Both ways of reasoning give the same result, since only those investment projects whose expected returns are greater than the current interest rates will be carried out. Then, the more interest rates increase, the more investment projects fall short compared to this financial norm. Because of credit market imperfections (Box 9.4),

BOX 9.4

CREDIT MARKET IMPERFECTIONS

Credit market imperfections stem from the fact that firms, as well as human beings, are not assumed to be spontaneously trustworthy, so that credit contracts may not be forthcoming. Information asymmetries between lenders and borrowers may generate suboptimal credit decisions. For example, banks do not perfectly know the financial soundness of borrowing firms, and they may end up with lending money to the firms that cheated the most to present a good picture, while the honest firms may have appeared too risky. While the previous example deals with adverse selection (pre-contract opportunism), another credit market imperfection induces moral hazard problems (post-contract opportunism): banks cannot always control what firms may do next, and firms may become less cautious, once the credit has been granted. To protect themselves from these behavioural failures that could endanger borrowers' ability to reimburse, banks are thus obliged to make borrowers pay a premium on the interest rate they would have paid in the absence of information asymmetries. Consequently, these heightened financing costs may discourage firms from borrowing and investing.

the shocks stemming from credit conditions may amplify the variations in investment, and thereby in the economy as a whole.

For New Keynesian economists such as Ben Bernanke (see Bernanke et al., 1999), this gives rise to a financial accelerator. If the economic situation appears riskier, banks may claim a higher interest rate premium and restrain credit. Firms' investment is therefore expected to drop, and with less (investment) spending, the economy slows down. In good times, the opposite is true: banks will lower their interest rate premium and deliver more credit, with an increase in firms' investment and an improving macroeconomic environment. In the end, banks' appreciation of risks is procyclical (it follows the state of confidence in the economy at large), but interest rate premia are countercyclical (they move in the opposite direction of the economy). Beyond banks' own evaluation of risks, the economy is also affected by monetary policies. When the central bank changes the policy rates of interest, commercial banks endure a change in their refinancing costs that is transferred into the interest rate they apply to borrowers. If the central bank decides to lower its interest rate (expansionary monetary policy), commercial banks will lower their own rates, which could lead to an increase in investment for firms, and to a boost of economic growth. The reverse is true when the central bank decides to increase its interest rate (restrictive monetary policy): commercial banks increase their own rates, credit and investment are reduced, the economy slows down. Consequently, it could be argued that macroeconomic cycles are crucially determined by monetary policies.

Finally, when accounting for the desire to invest by all firms, we are able to construct a macroeconomic investment function. In the general equilibrium framework, the investment function is reducible to a demand for capital on what we call the loanable funds market. On this market, the supply of loanable funds refers to savings, which are thought to be positively related with the interest rate. For example, households are considered to be prone to save a bigger share of their incomes when interest rates rise, as their savings would yield a higher return. Another way to justify this relationship between saving and the interest rate is to assume that an increase in the interest rate will incite households to postpone their consumption spending now so as to consume more in the future thanks to interest incomes.

For the reasons explained above, the demand for loanable funds is lower when the interest rate increases (opportunity cost or financial burden of indebtedness). The loanable funds market is thus the place where saving and investment are equalized through variations in the interest rate: the forces of supply and demand in the loanable funds market. Provided there is no obstacle to

BOX 9.5

THE CAMBRIDGE CAPITAL CONTROVERSY

During this controversy, economists from Cambridge (UK) faced economists from Cambridge (US) on the concept of capital. On the English side, Joan Robinson, Luigi Pasinetti and Piero Sraffa criticized the treatment of capital in the neoclassical theory defended by Paul Samuelson and Robert Solow on the American side. For Robinson, Pasinetti and Sraffa, it is impossible to aggregate heterogeneous forms of productive capital in a simple production factor, as the neoclassical production function does. After different contributions from each side, Samuelson finally admitted defeat, but most of the academic literature chose to ignore the end of this controversy and continue to use aggregate capital in macroeconomic models (eventually with the assumption that there is only one type of capital goods).

these variations, saving will always equal investment, thereby respecting Say's Law: as a result, there is no possibility for a general crisis of overproduction, since households' saving (the supply of loanable funds) is necessarily compensated by firms' investment (the demand for loanable funds), so that aggregate production will be sold either to households (consumption goods) or to firms (investment goods). Consequently, this reasoning assumes that production goods are homogeneous: in its simplest formulation, the theory here assumes that there is only one type of goods, but this product could serve different uses (namely, consumption and investment). However, this vision is proven wrong, as explained in the so-called 'Cambridge capital controversy' (Box 9.5).

Keynes's approach to investment

The approach of John Maynard Keynes to the theory of investment differs radically from the mainstream views. Indeed, Keynes does not support the same institutional features concerning the functioning of markets. More precisely, for Keynes, there is no such thing as a loanable funds market: investment is not brought into equality with saving as a result of variations in interest rates. As a consequence, Keynes (1936) opens the way for the possibility that desired saving may be greater than desired investment. If this occurs, the economy could suffer from insufficient demand, thus generating involuntary unemployment. If the interest rate is no longer determined by the confrontation between investment and saving on the loanable funds market, as in neoclassical economics, Keynes has to provide a new theory of interest rate determination. This is precisely the purpose of his theory of liquidity preference, where the interest rate is the price borrowers have to pay

lenders to make them give up their liquidity. Savers spontaneously prefer to hold their wealth in the form of bank deposits (perfectly liquid assets), but they could also keep their wealth, for example, in the form of bonds (less liquid than bank deposits) if their interest rate is high enough to compensate savers for the loss of liquidity (bank deposits). In the end, the interest rate depends on the amount of wealth and the reluctance of savers to hold illiquid assets.

While there is no place for the interest rate in Keynes's theory of investment in a neoclassical sense, this is not because there is no loanable funds market in Keynes's thought, nor because the interest rate does not equalize savings to investment. Rather, the interest rate impacts upon investment decisions, but it does so in a different way than in neoclassical economics. While neoclassical economics states that firms invest as long as the marginal productivity of capital is superior to the interest rate, Keynes's theory posits that firms invest as long as the marginal efficiency of capital is superior to the interest rate. For a rapid reader, the difference is not striking. The marginal productivity of capital designates the incremental profit stemming from the last unit of investment, while the marginal efficiency of capital stands for the expected profits for the last unit of investment. Once again, the distinction relies on a simple word: 'expected'; but this single word conveys a whole new theory. Indeed, for Keynes, the key variable is no longer the interest rate, but these expected profits, and this therefore places at the core of his approach the notion that the future determines what happens in the economy today. If firms are optimistic about their future sales, they will have a good appreciation of the return on new investment, and they will invest more even though the interest rate stays the same. Keynes emphasizes this psychological component in the evaluation of investment projects.

But there is a problem: in the long run it is very difficult to make predictions, and probability will not offer you a safe framework. Specifically, when dealing with investment decisions, firms cannot know in advance: (1) the effective durability of the new equipment (five or ten years, less or more?); (2) the demand for their products just a few years ahead (possibility of new competitors and/or new products on the market, variations in national income, and so on); (3) the structure of their costs (changing labour costs, changing fiscal environment, fluctuations in energy costs, and so on); and (4) the evolution of interest rates (which can make saving become more financially interesting than productive investment).

Several times, Keynes writes that we simply do not know what the future will resemble: we live in a world of what post-Keynesians call 'radical uncertainty', where: (1) we cannot know in advance the full list of possible future events;

BOX 9.6

RISK AND RADICAL UNCERTAINTY

While in the common language there is no clear distinction between risk and uncertainty, post-Keynesian economics highlights the essential difference between the two terms. A risky environment is a situation where one can know every possible issue, and where it is possible to count on a probability distribution for these different issues. As a consequence, one is able to make calculations. For example, in a classic dice game, there are only six issues for every toss, and every issue has the same probability to happen, provided the dice is not rigged. However, the economy is not as simple as a dice game, or even as a card game where one can expect to count cards so as to calculate probabilities and make good bets. A radically uncertain environment is a situation too complex to allow for such calculations. It is impossible to know what the exchange rate will be in five years: some totally unpredictable events may occur between now and then.

and (2) we cannot associate a distribution of probability to these different issues (Box 9.6).

Consequently, it is impossible to make decisions through the use of probabilistic calculations on the profitability of investment. What will enforce the decisions of an entrepreneur is not the exercise of these calculations, but the psychological impulse of acting instead of doing nothing. Keynes puts forward the concept of animal spirits to describe this instinctive dimension of economic decisions, and specifically for investment decisions. The improving of animal spirits would thus refer to the strengthening of optimism amongst firms, beyond the simple assessment of effective economic conditions. These animal spirits are obviously exposed to sheep-like behaviour: entrepreneurs observe one another, and the optimism as well as the pessimism of some of them may easily be contagious to others, so that the thirst for investment may be subject to waves of euphoria (causing booms) and sudden panic (generating crashes).

According to Keynes (1936), the changing mood is not confined to firms investing in productive equipment. Financial markets are also subject to bouts of irrational optimism and pessimism. The value of the firm on the stock exchange may undergo brutal variations, which in turn may have an impact on productive investment. Indeed, for Keynes, because of uncertainty, there is no reason why a firm would buy new equipment through an uncertain productive investment if it can acquire equivalent equipment by buying the firm on the stock exchange. The value of a specific firm on financial

markets thus reveals in some way the second-hand value of its capital stock, while the price of productive investment reflects the value of new equipment. If the value of a competing firm on the stock exchange is low, a firm wanting to increase its productive capacities may benefit from buying this competitor instead of buying new equipment.

The consequences of purchasing a firm on the stock exchange (a financial investment) or purchasing new machines (a productive investment) are manifold. At the microeconomic level, this kind of financial investment is certainly equivalent to productive investment. At the macroeconomic level, however, financial investment is not as good a solution as productive investment, because it does not add new productive capacities (therefore with no positive effect on economic growth and employment), and it also reduces competition (with a potential effect on income distribution and a decline in the wage share).

Reciprocally, in periods of high stock market prices, productive investments may be increasing: firms may be encouraged to buy new productive equipment, because buying a competitor is too expensive, but also because the good valuation of stocks on the market signals confidence in future profits. The high stock market prices are also the promise of important financing in the case of new equities issues. American economist James Tobin (1969), a Nobel laureate in economics in 1981 (Box 9.7), suggested a new concept to capture this potential positive effect of financial markets on productive investment: the now famous 'Tobin's Q', which is the ratio of the value of existing productive equipment (reflected through firm's valuation on the stock exchange) to the value of new productive equipment (on the equipment goods market)

If Tobin's Q is higher than one, it means that it is preferable to buy new productive equipment instead of buying a competitor to acquire existing

BOX 9.7

JAMES TOBIN

James Tobin (1918–2002) was a Keynesian economist who gave his name to a political proposal: the 'Tobin tax' is supposed to discourage financial speculation while taxing short-term capital movements. James Tobin received the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel in 1981, but it was not for this political proposal: the Nobel Prize was awarded to Tobin for his works on financial markets and their relations to spending decisions (most notably, the Q theory of investment).

productive equipment. When financial markets are on the rise, Tobin's Q increases, and this is considered to be favourable to productive investment.

Beside these positive aspects, what did Keynes think of the stock exchange market? As it turns out, Keynes also blames financial markets for being a poor guide to investment. The stock exchange is often the place for speculative forces that give too little attention to the effective productive projects behind firms' stocks. The stock exchange may first be seen as a positive institution for productive investment, as it offers financial investors the liquidity needed for their investment. On financial markets, the liquidity is the ability to sell one's asset at every moment in time. This liquidity allows timid investors to engage in financial investment, because they can get out at any time, whereas it could be difficult to depart from a productive investment, which is often irreversible: it is easier to sell the stock of a windmill company than to sell a windmill itself.

Theoretically, the stock market could thus kill two birds with one stone: (1) provide the necessary liquidity to financial investors; and thus (2) attract saving to financial markets so as to contribute to firms' financing. For purposes of clarification, only new equity issues bring financing to firms; most of the activities on financial markets are concerned with trading existing stocks, thus granting stockholders the liquidity for their assets. However, liquidity does not come without drawbacks: while financial investors may sell their stocks at any time, there are continuous revaluations of firms' stock prices. Considerable efforts are devoted to trying to guess stock prices in the very near future instead of finding out which investment may be profitable to the economy in the long run. In the end, the thirst for liquidity on financial markets may become an obstacle to productive investment. Indeed, as Keynes (1936, p. 159) reminds us:

> Speculators may do no harm as bubbles on a steady stream of enterprise. But the position is serious when enterprise becomes the bubble on a whirlpool of speculation. When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done.

When financial markets are organized in such a way that favours speculation, the fluctuations of stock prices may be detrimental to productive investment. To counter these destabilizing factors, Keynes was supportive of a tax on all transactions in an effort to reduce speculation. But he was also advocating for the state to undertake a significant share of total investment to stabilize the economy. This 'socialization of investment', as Keynes calls it, is seen to be a solution to the inherent instability of private investment, but also as a way to defeat the global scarcity of capital that leads to regular unemployment in

Keynes's thought. In post-Keynesian economics, investment is thus thought to be very volatile, especially in the context of an uncertain future.

Let us move now from the question of the determinants of investment to the question of its consequences. To do so, we have to focus on a given state of long-term expectations, that is, expectations of the very far future. In doing so, we will assume the following: we will consider (1) entrepreneurs' mood as a constant; and (2) investment as exogenously given, meaning we will assume that it does not depend on the interest rate or other variables, and it remains the same for our purpose. Given these assumptions, it is possible to briefly present the multiplier mechanism. This model starts with some (further) simplifying assumptions: (1) the economy is closed (no relationships with the rest of the world); (2) there is no public intervention (neither public spending nor taxes). In such a stylized economy, aggregate demand is limited to the sum of consumption spending by households and investment spending by firms. If we add the hypothesis that households receive both wages and the total amount of profits (say, through dividend distribution), it is possible to write down the equivalence between aggregate supply and households' income. The next step is that we assume households consume according to the following function:

$$C(Y) = C_0 + cY \tag{9.1}$$

where c is the marginal propensity to consume, Y is households' current income (we assume no taxes), and C_0 is the autonomous consumption. Aggregate demand then becomes:

$$Y_d = C(Y) + I \tag{9.2}$$

where I stands for firms' investment, which we assume is exogenously given, following the assumption of a constant level of long-term expectations.

With the standard equilibrium condition for the goods and services market that aggregate demand equals aggregate supply (Y_s), we can now find an expression for the equilibrium income (or equilibrium production) as follows:

$$Y_s = Y_d \tag{9.3}$$

Re-arranging, we find:

$$Y = C(Y) + I = cY + C_0 + I$$

$$(1 - c)Y = C_0 + I$$

$$Y_e = (C_0 + I) / (1 - c) \tag{9.4}$$

Provided the propensity to consume is lower than one, it is possible to show that an increase in investment (following an improvement in optimism on behalf of the firms) will lead to a bigger increase in equilibrium income. That is precisely what is called the investment multiplier. Mathematically, this multiplier (k) is simply:

$$k = 1 / (1 - c) \ \ (\text{with } k > 1 \text{ if } 0 < c < 1) \tag{9.5}$$

The higher the propensity to consume, the higher the value of the multiplier, k. This stands to reason: if we all increase our consumption via a greater propensity to consume, we create more spending.

The implications are very important: at the macroeconomic level, a small variation in investment may generate a huge variation in economic activity. This result is important for understanding how public policies could bring back full employment, since a small stimulation of investment spending could contribute to the creation of many jobs. But one should also recall that investment is subject to important variations, so that a wave of pessimism may provoke a marked drop in investment, which could lead to an economic collapse due to the investment multiplier. Therefore, public intervention is all the more needed to stabilize the economy.

Graphically, the multiplier principle could be illustrated through the 45° diagram. On the horizontal axis, we reproduce aggregate supply. On the vertical axis, we have aggregate demand and its components. To confront aggregate supply and aggregate demand, we need to project aggregate supply on the vertical axis, and the bisector (the 45° straight line, which gives its name to the diagram) should be drawn precisely for this reason. The equilibrium condition is then materialized by the intersection between the demand schedule and the bisector. To take an example, we can go back to the previous consumption function and specify that $c = 0.75$, $C_0 = 5$, and to fix $I_1 = 5$. These numerical simulations give the diagram illustrated in Figure 9.1.

In Figure 9.1 equilibrium is reached when households income amounts to 40. Then, we can look at the evolution in this very simple economy when an optimistic wave makes investment rise to $I_2 = 10$. Graphically, it implies that the intercept of the aggregate demand schedule (Y_d) is moving upwards: even when aggregate income is zero, there is demand thanks to autonomous consumption and investment, but now investment has increased from 5 to 10, making the intercept increase from 10 to 15. Since the propensity to consume has not

Figure 9.1 The 45° diagram

Figure 9.2 The multiplier principle inside the 45° diagram

evolved, the slope of the aggregate demand schedule remains the same as before, and the same as the consumption schedule, that is *c* itself: indeed, the slope of the aggregate demand schedule indicates the increase in demand coming from an increase of one unit in aggregate income, and the propensity to consume accounts precisely for the share of an increase in income that is dedicated to consumption. According to the multiplier principle, this increase of investment should generate a bigger increase in equilibrium income (Figure 9.2).

Graphically, we observe that the equilibrium is actually moving up to a production level of 60. A variation in investment of 5 units leads to a variation in income of 20 units. The reason behind this multiplier principle is that the initial increase in investment of 5 units creates an excess demand on the goods market. Consequently, firms will try to produce more to satisfy

this demand. But, while increasing their production, firms also increase the distribution of wages and profits towards households. This increased distribution of incomes incites households to spend even more on consumption. Firms thus face a higher demand for their products, so that their first increase in production to address the initial situation of excess demand is insufficient. This second round of excess demand pushes firms to increase again production once more. The same cause generates the same consequence: firms can only produce more while paying more wages and/or profits to households, who will increase their spending, thereby inducing firms to raise production again, and so on. This virtuous circle is not without limits, since households do not consume the totality of the additional income at each stage of the reasoning (this is precisely the meaning of the assumption $c < 1$). As time goes by, firms are finally able to scale back aggregate demand, but this implies in the end producing a lot more than the initial increase in investment.

The investment multiplier is surely one of the most famous of Keynes's ideas. It recognizes the consequences of variations in investment. Yet, Keynesian economics has also developed theories on the determinants of investment. The accelerator principle represents the most significant contribution in this regard. This principle is older than Keynes's (1936) *General Theory*. Its first account appears in French in the writings of Albert Aftalion (1908), but John M. Clark (1917) formulated it explicitly later (Box 9.8). The basic idea is that

BOX 9.8

ALBERT AFTALION AND THE METAPHOR OF THE STOVE

Before the accelerator model of John M. Clark (1917) and the oscillator model of Paul Samuelson (1939), the French economist Albert Aftalion underlined the accelerator mechanism in 1908. He used the metaphor of a stove to account for the economic cycles generated by investment. In a too-cold room, people tend to overload the stove. But, as coal burns slowly, the heat increases progressively, and coal's overload ends up pushing the temperature too high. People are then forced to open the windows so as to refresh the room, potentially up to the initial point when the room was too cold. Applied to economics, this metaphor helps us to understand the time discrepancy between the decision to invest and the result of investment: firms decide to invest on the basis of the actual and expected demand, but their investment will contribute to increase this demand, which incites them to invest even more, so that they finally overinvest. To eliminate excess capacity, firms are then obliged to cut investment drastically. As a consequence, the economy declines and excess capacity is removed, until firms decide again to invest because of expected demand.

BOX 9.9

PUTTY OR CLAY?

Production techniques are called 'putty' when firms can adjust their choices of production factors, switching from a capital-intensive technique to a more labour-intensive one, depending on the evolution of factor prices. This flexibility of production techniques requires the substitutability of production factors: one can replace capital by labour, or conversely. When production techniques are meant to be 'fixed in clay', production factors are complementary: one cannot use more machines without more workers, since labour and capital are used in fixed proportions. Neoclassical economists support 'putty' models while heterodox economists are more used with 'clay' models. Nevertheless, because of the potential irreversibility of initial investment, some theories now distinguish between 'putty–putty' models and 'putty–clay' models: at the creation of a new plant, the choice is open between labour-intensive or capital-intensive production techniques ('putty'), but then, according to the industry, either it may be impossible to change this initial choice ('putty–clay'), or the production techniques remain relatively reversible because of the absence of specificity in capital goods ('putty–putty').

investment is now made dependent on the expectations of future demand, rather than merely being considered exogenously determined, and the implications are important.

Contrary to neoclassical economics (which assumes substitutable factors of production), the accelerator principle relies on the notion of complementary factors. Concretely, this means that firms cannot increase production while keeping the same capital stock and hiring more workers. The Keynesian literature says that production techniques are 'fixed in clay', while they are 'putty' in neoclassical economics (Box 9.9).

Since firms cannot modify their production techniques, the relationship between production and capital stock is straightforward: according to the level of demand one expects or anticipates, one can simply calculate the capital stock necessary to produce the adequate quantity of goods and services (K^d), multiplying this expected demand (Y^e) by the fixed value for the technical coefficient of production (ν, defined as the ratio between the capital stock and production). Hence we get:

$$K^d = \nu Y^e \tag{9.6}$$

Since investment, net of depreciation, is by definition the variation of the capital stock ($I = \Delta K$), it comes out that investment is determined according to the following formula:

$$I = v\Delta Y^e \tag{9.7}$$

What is important here is that investment is not determined by the level of expected demand, but by the variation of expected demand. Provided that the technical coefficient of production (which is also here the accelerator coefficient) is higher than one, investment will be extremely volatile, with large amplifications in variations in demand. Now, in the real world, some conditions reduce the value of this accelerator coefficient: firms will not necessarily increase production in response to the increase in demand: (1) if they have inventories; (2) if they perceive the increase in demand to be only temporary; or (3) if they choose to raise their prices instead. Moreover, firms will not necessarily invest to answer the increased demand: (1) if their productive capacities are not fully used; (2) if the investment goods sector is unable to satisfy the supplementary demand due to bottlenecks; or (3) if they are unable to finance these new investment projects.

In the simplest model, expected demand is estimated based on current demand, but there are more complex versions of the accelerator principle, where expected demand is measured through a partial adjustment to past demand, with a declining weight for old demand (Koyck, 1954). This flexible accelerator model allows one to take into account some of the lags implied by the conditions explained above.

When combining the accelerator principle and the multiplier principle, it is possible to arrive at the so-called oscillator model stemming from Paul Samuelson (1939). This model tries to present the evolution of an economy where investment depends on demand, with amplified responses of investment to variations in expected demand (accelerator principle), but the model also includes the fact that variations in investment lead to strong variations in production (multiplier principle). This model is highly sensitive to the assumptions placed on crucial parameters: depending on the respective values for the propensity to consume (which determines the size of the multiplier) and for the technical coefficient of production (which determines the speed of the accelerator), the model may exhibit diverging, converging or even regular cyclical paths for the economy. The diverging cyclical path describes an economy gravitating towards an equilibrium from which it goes further at every rotation: the economy alternates phases of booming growth and collapsing crises where ups and downs are ever larger as time goes by. The converging

Figure 9.3 The oscillator model

cyclical path refers to the fact that the economy undergoes ever-cushioned fluctuations, with ups and downs progressively disappearing with an economy finally reaching a stationary state. The regular cyclical path again represents fluctuations around a precise point, but this time the scale of booms and busts is constant over time. The oscillator model may reproduce very different dynamics according to the calibration of the key parameters (Figure 9.3).

The calibration process indicates the choice of value given to the different parameters entering the model's equations. A big part of the job of an economist is precisely to discriminate between different values for a given parameter, in order to represent the evolution of the economy: the parameter value has to be relevant, and then the model may become a good abstraction of a complex reality.

To take an example, consider the simple following functions with different calibrations as indicated below:

$$Y_t = C_t + I_t$$

$$C_t = cY_{t-1} + C_0$$

$$I_t = v(C_t - C_{t-1})$$

The first equation means that aggregate production is composed of consumption goods and investment goods. The second equation describes the aggregate behaviour of consumers: they consume according to their previous period income (the subscript t refers to the time period, so that t–1 stands for the previous period), but the consumption function also includes an autonomous component (C_0). Concretely, it could mean that households receive their wages at the end of the month, and they spend them in the following month. The third equation describes the aggregate behaviour of firms regarding investment. Equations (9.6) and (9.7) are only slightly modified: firms have a desired capital stock depending on their expectations of demand and on the available technologies, which fixes a certain proportion between the capital stock and production (v); we suppose here that firms' expectations on the variations of demand are based on the past variations of consumption, so that firms determine their investment spending according to the variation in consumption between the present and the last period, with the speed of adjustment being here the technical coefficient of production (v).

The different values given to the parameters entering the model's equations produce here three different dynamics with increasing, decreasing or constant oscillations. In the model, the cycles thus appear to be unavoidable, with crises always coming back at regular time periods. If this property of the simulated economy also applies to the real world, it implies that a simple behavioural change may generate deep transformations because of retroactive processes between investment and demand. For example, the difference between the left-hand figure and the middle one is based only on the value of the technical coefficient of production going from 1.5 to 1.3: the first case implies cycles with ever-growing booms and busts, while the second case generates ever-softened fluctuations. If we consider this model as being a relevant description of our real-world economies, knowing the precise value of this parameter should thus be a very important point for economic policy analysis, since the urgency as well as the type of public intervention may not necessarily be the same in the two different contexts. Hopefully, real-world economies seem more resilient than these simple cases: cycles in the real world do not seem to come back at fixed time period, nor to be ever-increasing, decreasing or regular. If cycles do seem to exist in the real world, they do not appear to be so 'natural' as is the case in the oscillator model. It is possible

to avoid crises for a considerable period of time, but then crises may surge repeatedly: while there was no major crisis during what has been called the Golden Age of capitalism (1945–70), there has been an outbreak of crises since the beginning of the 1980s (the monetary shock of 1981, the financial crash of 1987, the European crises of the early 1990s, the Internet bubble burst in 2001, the consequences of the subprime crisis since 2007).

The simple oscillator model does not seem to offer a relevant description of our real-world economies. A first reason is that this model is simplistic and includes unrealistic assumptions. It is simplistic because investment is determined only by naive expectations of future demand based on past demand. It is unrealistic because the model excludes the relations with the rest of the world or public spending. A second reason why there is such a distance between the real-world economies and the predicted model is precisely that the state may succeed in stabilizing an inherently unstable economy. This is maybe the main achievement of Keynes's inspired economic policies: because of the variability of firms' investment decisions, and because of the amplified impact of these decisions on economic activity, it is all the more important for the state to smooth private investment fluctuations through countercyclical macroeconomic policies. Concretely, it means that when firms stop investing, governments should spend more in order to avoid a complete economic collapse. Conversely, when firms are involved in an investment fever, it is also important that governments avoid the overheating of the economy while reducing their spending. Countercyclical macroeconomic policies imply that governments take the opposite stance of firms, which means that public deficits are expected to rise during a crisis: this is the price to be paid to balance economic cycles. The Golden Age of capitalism is often associated to a period of Keynesian macroeconomic policies. Under this doctrine, economic crises were scarce. In contrast, since the beginning of the 1980s, the influence of Keynesian ideas in public policies has undergone an important downturn, and the frequency of crises has risen significantly.

The heterodox perspective

The previous sections discussed the neoclassical investment model as well as Keynes's and Keynesians' views of investment. This section turns to the views of post-Keynesian economists, or generally the heterodox school.

The heterodox perspective on investment can be complicated to understand, as it encompasses various strands. Some authors insist on the role of profitability in determining investment decisions, while other authors prefer to

stress other factors (such as demand on the product market, expectations, financing, institutions). It is impossible in this chapter to present every single approach to investment in the heterodox tradition, and this section will focus on some of the main contributions of the post-Keynesian and institutionalist schools of thought.

Henry Roy Forbes Harrod is surely one of the best figures of investment (see the portrait of Harrod at the end of this chapter). His work contributed to the development of a productive literature on the question of instability. The starting point of this analysis is the recognition of the ambivalence of investment: on the one hand, investment is a way for firms to add new equipment to their capital stock, and thus to increase and/or improve their productive capacities (the capacity effect); on the other hand, investment is also a spending that is intended for producers of equipment goods, and as such it is a component of aggregate demand (the demand effect). The interplay between these two effects may contribute to destabilize the economy, despite rational adjustments by firms.

First, let us assume that firms desire to operate at a target utilization rate for their productive capacity. Let us fix this target at 85 per cent, which roughly reflects current practices for some industries. This means that, on average, firms will not produce at full capacity, but at a level less than full, which reflects a 'normal' level of capacity utilization. In other words, this assumption suggests that it is costly for firms to own unused productive equipment, but it is also risky to operate at full capacity in case of an unexpected surge in demand on the market for produced goods and services. Assume the actual utilization rate is at 75 per cent: firms operate below their target. This means that demand on the product market (featuring at the numerator of the utilization rate) occupies only 75 per cent of firms' productive capacity (denominator of the utilization rate): firms have idle productive capacities. This can be seen as 'elbow room' in case there is an unexpected increase in demand. The rational behaviour is thus to invest less than the expected growth rate of demand, so that the increase in demand on the product market will be partly satisfied by the use of the previous productive capacities' reserve. Respectively, in the case of an initial actual utilization rate of 95 per cent, the rational strategy for firms is to increase investment faster than the expected growth rate of demand, so as to accumulate much new equipment and restore the target utilization rate of 85 per cent.

For now, our reasoning has focused only on the capacity effect of investment. The problem is that the reaction of firms to the disequilibrium of their utilization rate will widen the gap between the actual and target rates

of utilization, instead of filling it. This is so because of the demand effect of investment. It requires a clear explanation.

The slowdown of investment in the initial case of underutilization will contribute to a slowdown of productive capacities, but will also reduce aggregate demand on the product market, because less investment means less demand for producers of equipment goods, but also because of the multiplier effect of changes in investment. Globally, this demand effect may overcompensate the capacity effect: the impact of variations in investment is bigger on the numerator of the utilization rate than on its denominator. This means that either the economy is initially at equilibrium, with firms operating at the target rate of utilization, or the economy will continuously move away from it, despite firms' rational adjustment at the microeconomic level.

For Harrod (1939), long-term economic growth rests on a knife-edge: either you are on the good track, or you will fall apart. The rate of capital accumulation (that is, the ratio between investment and capital stock) has to be equal to what Harrod called the warranted rate of growth: this is the only rate of accumulation compatible with the dynamic preservation of the goods market equilibrium between investment (I) and savings (S). The short-run equilibrium condition on the goods market is simply extended to the long run: the equality $I = S$ becomes the equality between the rate of accumulation (I / K, where K is the capital stock) and the savings rate (S / K), which can be rewritten as:

$$g^s = (S / Y) * (Y / K) = s / v \tag{9.8}$$

The long-run equilibrium condition is therefore achieved when the rate of accumulation is equal to $g^i = g^w = s / v$ (where g^w is the warranted growth rate). The obvious problem is that there is no reason why firms will spontaneously decide to invest at such a speed. Therefore, it is very likely that the economy will be permanently out of equilibrium in the long run.

A second problem for Harrod is that, even if this peculiar condition is met, the economy is not necessarily at full employment. For now, we only allude to the goods market situation. In order to include the question of unemployment, we need to bring back into this picture the natural rate of growth, which corresponds to the sum of labour force growth and productivity growth ($g^n = n + \mu$). In order not to see employment ever decreasing or increasing, the rate of accumulation that equilibrates the goods market ($g^i = g^w$) has to be equal to the natural growth rate ($g^i = g^w = g^n$). Again, there is little chance that firms will decide to invest at this precise rate. For Harrod, the economy is thus prone to bouts of instability because of the dual nature of investment.

Figure 9.4 The banana diagram

In the post-Keynesian school, the theory of investment has focused not only on demand, but also on the links between investment and profit. For example, in her so-called banana diagram (Figure 9.4), Joan Robinson (1962) proposed a model with feedback effects between investment and profit. On the one hand, present-day investment is influenced by past profits for two different reasons: (1) past profits may be regarded as a good indicator of future profitability; but (2) past profits may also induce investment because of the financing they allow. On the other hand, present profits are determined by present investment. This second relationships is deterministic, while the first one is more on the incentive mode: 'the rate of profit [is] a function of the rate of accumulation that generates it', while 'the rate of accumulation [is] a function of the rate of profit that induces it' (Robinson, 1962, p. 48). This determination of profit by investment comes from what is called Kalecki's profit equation (see Box 9.10; and Chapter 12 for a portrait of Kalecki).

To give an account of it, let us assume that we are in an economy with neither public intervention nor international relationships. Also assume that workers spend the totality of their income (marginal propensity to consume is equal to 1). As a consequence, the only savings in this economy come from firms, which save a share (s) of their profits (Π), and the equilibrium condition between savings and investment is:

$$s\Pi = I \tag{9.9}$$

BOX 9.10

KALECKI'S PROFIT EQUATION

For Michał Kalecki, in a simplified economy, profits are determined by investment spending at the macroeconomic level. This result is obviously not valid at the microeconomic level: it is not because a certain firm decides to invest a lot that it will gain much in profits. To demonstrate this result, we just have to begin by simplifying assumptions: there is neither public intervention, nor trade with the rest of the world. Thanks to these simplifying hypotheses, we know that output (Y) is demanded only through consumption spending (C) and investment spending (I): $Y = C + I$. If we add the assumption that workers do not save ($S_w = 0$), workers spend their entire wages income (W) in consumption ($W = C$). Then, we just have to recall that national income (Y) is simply the sum of wages and profits (Π). Finally, we get:

$$Y = C + I = W + \Pi \leftrightarrow \Pi = I$$

This equation is the reduced form of Kalecki's profit equation. Kalecki states that the direction of causality goes from investment to profits, because capitalists can decide the amount of their investment spending, but cannot decide the amount of their profits. Firms manage to make profits when they succeed in receiving more in turnover than they engage in costs, and investment spending is precisely a flow that increases firms' turnover at the macroeconomic level, while it does not correspond to a cost. Therefore, it is possible to extend this logic while relaxing simplifying assumptions. To sum up Kalecki's profit equation, the famous aphorism is often used: 'capitalists earn what they spend, while workers spend what they earn'. But capitalists do not spend only in investment: capitalists' consumption thanks to dividend income (C_Π) is also a macroeconomic part of their turnover, which is not counted as a cost. Public deficits also participate to profits, since public spending (G) increases firms' turnover more than taxes (T) are part of firms' costs. Trade surpluses are also a determinant of macroeconomic profits, because exports (X) allow resident firms to capture a source of demand they do not initiate through their distribution of wages. Reciprocally, imports (M) constitute a loss of profits for resident firms, because they distribute wages that do not reflux in turnover. Finally, workers' saving is obviously a loss of profits for firms, because it represents a share of distributed wages that again does not reflux. The extended profit equation is:

$$\Pi = I + C_\Pi + (G - T) + (X - M) - S_w$$

To consider the long run, we have to look at saving and investment rates rather than levels: it is possible to do so while dividing both ends of the previous equation by the value of capital stock. After just a little manipulation, it yields what is known as the Cambridge equation, where the macroeconomic rate of profit (r) is determined by the rate of capital accumulation (g^i) and the propensity to save (s):

$$r = g^i / s \tag{9.10}$$

The relation is the first element of Robinson's banana diagram. The second element is an investment function where investment is positively influenced by profits. In fact, for Robinson, firms require more and more profits for a given increase in capital accumulation. This non-linear relation when represented in the same diagram gives birth to the banana diagram in Figure 9.4.

These two relations generate two equilibrium points (A and B), where the profit rate resulting from firms' investment is just equal to what is required by firms. Now, while the low equilibrium (A) is unstable, the high equilibrium (B) is stable. To show this, assume firms have undertaken the accumulation rate g_c: we can see that they will benefit from a higher profit (C) than expected (D), so that we can expect them to become more optimistic while increasing their investment until the high equilibrium is reached (B). On the contrary, the low equilibrium (A) is unstable: either the economy is initially there, or it will necessarily diverge from it through a complete collapse (if the initial accumulation rate is just a little lower than g_a) or through a convergence towards the high equilibrium (if the initial accumulation rate is just above g_a). Finally, this banana diagram also figures out one of the main results in Keynesian macroeconomics, namely the paradox of thrift. An increase in the propensity to save (s) will indeed flatten the Cambridge equation of profit determination, so that the high equilibrium will move down, with a reduced accumulation rate and a diminished profit rate. In the end, despite the increase in the propensity to save, the amount of savings in the economy may be reduced because of this increased propensity applying to a decreased amount of profits.

After the emphasis on demand and capacity utilization in Harrod's writings, and after highlighting the double-sided relationships with profits in Robinson's banana diagram, the last strand of post-Keynesian literature surveyed here deals with the question of financing for investment (see Chapter 5 on the monetary circuit and the role of banks). There are plenty of authors on this line of research, and I will reference only some of them. In *The General Theory*, Keynes himself alluded to what he called borrowers' risks and lenders' risks, to distinguish the reluctance of firms to borrow and of banks to lend. Combined, these point to a possible limit in the volume of investment firms can (or desire to) undertake. Michał Kalecki develops the principle of increasing risk to describe the interplay between interest rates and indebtedness (Kalecki, 1937). The more indebted a firm, the higher the interest rate it has to pay. This principle of increasing risk is thus a potential constraint for highly indebted firms trying to invest. Hyman Minsky (see Chapter 6

for a portrait of Minsky) is also to be mentioned here, since he elaborated a theory on the cyclicality of the measurement of risk. When there has been no financial crisis for a while, banks and firms, having short memories, begin to neglect risks, and the cautious actions to slow the increase in indebtedness break down. Banks are likely to lend more and more to firms that are not necessarily financially sound; and firms themselves are increasingly prone to regard high level of indebtedness as perfectly safe. As a consequence, investments are going up as well as risks. Hence, the absence of financial crisis thus generates the risks that can contribute to a new financial crisis. This is what is called the paradox of tranquillity, often summed up as 'stability is destabilizing' (Minsky, 1975).

These first contributions of post-Keynesian authors on the financing of investment assemble to form a comprehensive study of firms' investment decisions at the microeconomic level. In the 1970s, a growing literature quickly established the foundation of the post-Keynesian theory of the firm. Through different authors (see Eichner, 1976; Harcourt and Kenyon, 1976), investment was usually analysed with respect to financing constraints. In what follows, I will present the theory of the firm as developed by Adrian Wood (Box 9.11).

In a 1975 book, Wood argued that the investment decision of firms is made at the intersection of two constraints – a financing constraint and a competitive constraint – both having implications on profit margins. Investment decisions are thus connected to pricing decisions. In a long-term perspective, firms try to achieve their objectives (in terms of profitability and market share) through a difficult balance between the need for high margins

BOX 9.11

ADRIAN WOOD

Adrian Wood is a British economist who laid the basis of the post-Keynesian theory of the firm in his 1975 book entitled *A Theory of Profits*. His education has notably been gained in Cambridge (UK), which shows the resilience of Keynesian ideas there. In his preface, Wood acknowledged many Keynesian authors who commented on earlier drafts of the book. Amongst them can be mentioned Wynne Godley, Geoffrey Harcourt, Richard Kahn, Nicholas Kaldor and Joan Robinson. With less echoes inside the post-Keynesian school of thought, Adrian Wood also contributed to the study of international development, showing for example that international trade may harm unskilled workers.

Figure 9.5 The post-Keynesian theory of the firm

in order to finance investment, and the need for low margins in order to capture demand. Even though Wood's original idea was about the relationship between accumulation and profit margins, the key structure of the post-Keynesian theory of the firm can be found in Lavoie's 1992 book, where the model can be pedagogically sketched out within a simple two-curve diagram that provides a relationship between profit rates and accumulation rates.

The first component of the traditional theory of the firm is the finance constraint, represented graphically by the finance frontier in Figure 9.5, which suggests that profits are a prerequisite for a firm wanting to invest. Profits are needed because they are a way to finance investment from within the firm, and are often referred to as internal finance. At the same time, banks will see profits as a sign of the firm's creditworthiness, and a profitable firm will also find it easier to raise funds by issuing new equities (external finance). The more profits a firm makes, the more investment it will be able to undertake. The finance frontier therefore represents the minimum profit rate a firm has to reach in order to secure the financing of its investment, given its ability to raise external funds and its commitment to pay dividends.

The second component of the traditional theory of the firm is the competitive constraint. This time, this constraint represents the maximum profit rate a firm can expect for a given rate of accumulation. For Lavoie (1992), there is a concave relationship between accumulation and profit expectations:

investment may first allow the firm to improve its position relative to its competitors (improved quality for the products). But then, if a firm wants to grow faster, it has to face both increasing selling costs (such as the cost of advertising) and increasing pressures to cut prices, so that the profit margins tend to decline in order to allow for an ever-growing market share. Inside this two-frontier framework, the firm will invest as much as permitted by the intersection of the finance and competitive frontiers (Figure 9.5).

The post-Keynesian theory of the firm is set within the context of what is called 'managerial capitalism', where firms are governed by managers whose objective is to maximize the growth of the firm, even though this may come at the expense of profits. For Galbraith (1967), these managers form a techno structure inside the firm. Managers do aim to grow the firm, because it increases their personal power, allows them to acquire specific interests, or even to build an industrial empire over which to reign. But, beyond these individual managerial dispositions for growth, the pro-investment bias also stems from the type of firms under study: post-Keynesians are mainly interested in the behaviour of big firms, what Eichner (1976) called 'megacorps' (see Chapter 7 for a portrait of Eichner). These big firms invest as much as they can, because they want to secure their survival in the long run, and growth is a way to gain power relative to competitors, suppliers or even the state: the bigger the firm, the easier it is to acquire competitors, to impose price cuts to suppliers, and/or to influence public policies and industrial regulations.

Yet, investment decisions do not only depend on demand and financing considerations, but also involve institutional factors. In a similar vein, Crotty (2003) deals with a regime of coercive competition where firms invest despite the lack of demand for their products: in some industries, firms may continue to invest despite low demand in the hope of surviving and being among the few firms that stand after their competitors are washed away by the downgraded demand environment. This overinvestment thus leads to chronic overcapacity in some industries. Of course, if firms invest despite the lack of demand and/or the absence of healthy financial conditions, the opposite is also possible: firms may refuse to invest in a favourable economic context. For example, Kalecki (1943) alludes to the potentially negative effects of full employment for capitalists. Indeed, with the disappearance of the threat of unemployment, workers are no longer disciplined in their claims concerning wage increases or improvements in working conditions. As a consequence, firms may face a surge in costs, and thus a decline in their profit margins. In order not to arrive at this gloomy perspective, firms may reduce their investment to prevent full employment. As a result, the lack

BOX 9.12

GOODWIN'S CYCLES

Richard Goodwin investigated the question of endogenous cycles in a Marxist tradition. Inspired by the dynamics of prey and predator populations, he proposed an application of this principle to the economic system. In biology, the prey population decreases when predators are numerous. But if predators are too numerous, the disappearing of prey may generate their own decline. When the predator population has dropped enough, the prey population starts to increase again, which allows for the return of predators. For Goodwin (1967), a similar logic may generate cycles in the economic life of capitalists and workers. If wages are positively related to activity and if investment is positively related to the profit share, then it is possible to highlight endogenously determined cycles: at first, when the level of economic activity is low, workers are not powerful enough to receive high wages, and the income distribution is advantageous to profits. Good profitability may incite firms to invest, which stimulates economic growth and employment, and thereby contributes to the increase in wages during this phase of expansion. When wage increases exceed productivity growth, the profit share begins to decline, which may incite firms to reduce investment, thereby precipitating recession, unemployment and wage contraction. When the restoration of profits is there, the time comes again to restart investment in a new cycle.

of investment may become a strong weapon to preserve capitalists' power regarding class conflict with workers. For some more radical theorists such as Stephen Marglin and Amit Bhaduri (1990), investment is thus directly linked to profitability, and a profit squeeze may be the cause of a generalized slowdown in both capital accumulation and economic growth (Box 9.12).

Empirical testing of investment functions

So far, I have listed several potential factors that determine firms' investment decisions. While orthodox economists insist on the role of factor costs (wages and interests), heterodox economists point at the importance of demand and financing constraints. Beyond the theoretical disputes between mainstream and heterodox economists, a full body of empirical evidence has emerged on the question of the determinants to investment.

A first set of empirical studies tries to measure the elasticity of factor substitution in response to factor prices. Recall that in neoclassical economics production factors are supposed to be substitutable, while they are complementary in heterodox economics. In neoclassical economics, therefore, a change in factor prices should cause a change in production techniques (to

more labour-intensive or more capital-intensive techniques). Concretely, an increase in wages relative to interest rates will incite firms to replace workers with new equipment. As a consequence, to support the neoclassical theory, empirical studies should find that the elasticity of factor substitution is close to unity (meaning that an increase of 1 per cent in the relative cost of capital respective to labour should entail a decrease of nearly 1 per cent of the capital-to-labour ratio). On the other hand, accelerator models of investment rely on the assumption of complementary factors and fixed production techniques. Mathematically, this implies that the elasticity of factor substitution should be equal to zero (meaning that when there is a change in the relative cost of capital and labour, there is no modification of the capital-to-labour ratio). If the literature does not yield undisputed measures for this concept, '[m]ost results seemed to suggest that the elasticity of substitution is closer to zero than to one' (Baddeley, 2002, p. 371). The influence of output demand on investment seems to be more robust than the neoclassical story about factor substitution.

Within heterodox economics, much empirical work tries to establish the strongest factor influencing investment decisions. While radical authors place the emphasis on profitability, post-Keynesian scholars prefer to insist on the role of demand (Box 9.13).

For radical authors, investment is positively associated with expected profit. Post-Keynesians also assume a positive relation, but the reasoning underlines the role of profit in financing investment. Both factors (profitability and demand) have always performed relatively well to empirical tests. But a burgeoning literature has also developed around the question of financialization (see Chapter 18 for a discussion of financialization).

Recall how, in the previous section, the post-Keynesian theory of the firm was presented as an integral part of an era of managerial capitalism. Since the beginning of the 1980s, however, financial liberalization has strongly modified firms' financial constraint: stockholders force firms to become more selective in their investment projects, to implement only those delivering a very high profit rate; what has been called 'short-termism'. Firms are also required to give up peripheral activities to external suppliers. The time of diversified empires is over, and stockholders demand that firms' managers concentrate on core activities, obviously the most profitable ones. The new doctrine imposed by stockholders also concerns profit utilization, with increased pressures to distribute profit through dividend payments and stock buybacks. As a consequence, in Figure 9.5, a given rate of accumulation now requires a higher profit rate than before: the finance fron-

BOX 9.13

POST-KEYNESIAN AND RADICAL VIEWS ON THE ACCOUNTING DECOMPOSITION OF THE PROFIT RATE

Regarding investment theories, post-Keynesian and radical explanations differ mostly on the question of identifying the main factor behind investment. If post-Keynesians insist on the role of demand (where the investment function depends positively on the utilization rate), radical authors prefer to underline the role of profits (with an investment function depending positively on the profit rate). These two views can be partly reconciled, most notably because demand is an important source of profit realization. Indeed, on an accounting basis, the profit rate is the ratio between the amount of profits (Π) and the capital stock (K), which can be decomposed as:

$$r = \Pi / K = (\Pi / Y)\,(Y / Y_{fc})\,(Y_{fc} / K) = \pi u / v \qquad (9.11)$$

While the ratio between aggregate output (Y) and full capacity output (Y_{fc}) is the definition of the utilization rate of productive capacities, the ratio between full capacity output (Y_{fc}) and the capital stock (K) defines the technical coefficient of production (v). In the end, the profit rate may increase because demand is increasing (u goes up) and/or because firms succeed in imposing a higher profit margin (π increases) and/or because firms are able to produce while using more capital and less workers (v declines).

An investment function where capital accumulation depends on both utilization and profit rates is thus about to introduce a strong demand effect: a direct one through the utilization rate, and an indirect one through the profit rate (which depends on the utilization rate). As a consequence, Bhaduri and Marglin (1990) suggested a kind of compromise function for post-Keynesian and radical authors where investment depends positively on two arguments: the utilization rate (u) and the profit margin (π). Post-Keynesian authors preserve the demand effect thanks to the presence of the utilization rate, and radical authors find an exploitation effect due to the presence of the profit margin. This investment function opens the possibility of nuanced story where, despite high demand, investment is not so strong because of low profit margins (profit squeeze theory).

tier moves upwards and investment is expected to be reduced because of financialization (Dallery, 2009). Nevertheless, financialization also presents firms with some opportunities. If we consider Tobin's Q theory, discussed earlier, the higher value of firms on the stock market may incite them to invest, because of the availability of new funds through new equity issuances. Financialization may also foster investment if firms earn more profits thanks to financial investment whose incomes may help relax the financial constraint. But as far as financial accumulation is concerned, it is also

possible to conceive of it as a substitute to productive capital accumulation, so that the total effect is a priori ambiguous. As for the question of the elasticity of factor substitution, the empirical literature is not unanimously conclusive. Yet, most studies seem to arrive at an overall negative effect of financialization on investment.

Concluding remarks

Since the subprime crisis, investment has been very low. Some people have interpreted this decline in investment as a new normal in a world entering a secular stagnation (see, for example, Summers, 2015 and his rediscovery of Alvin Hansen's work in the 1930s). For some economists (see Gordon, 2012), this situation may be caused by a faltering of innovation. More specifically, entrepreneurs have lost sight of what could be our economic future: while the society brought about by cars supplied several investment opportunities since the 1960s, firms lack an equivalent promise of great products, so that investment is reduced. At the end of the 1990s, the new technologies of information and communication have contributed to a massive investment boom. For some time, firms were involved in an investment fever, where everything related to the Internet or computers may be seen as profitable. This wave of optimism also triggered financial speculation, with the dot-com bubble bursting in 2001. Since then, despite the changes in our day-to-day life brought by new technologies, investment in the economy has returned to historical low levels and, for now, firms' imagination seems to have been only temporarily stimulated by new technologies.

Other explanations for the slowdown in capital accumulation point to the persistent low demand on the product market in the post-subprime crisis era: in spite of very low interest rates, firms are reluctant to invest because they suffer from insufficient demand for their products. Especially, growing inequalities may have contributed to the low-demand environment. Indeed, the wealthiest households are supposed to consume a lower share of their incomes and, as national income increasingly concentrates in their hands in many countries around the world, aggregate consumption is not as dynamic as it could be with a more equal distribution of income. For low-income households, an important engine for consumption has been credit instead of stagnant wages. But, following the financial crisis where banks were stressed, consumers' credit is no longer as available as before.

Finally, a whole literature blames financialization for being a cause of declining investment, beyond the recent slowdown provoked by the subprime crisis. Whatever the precise reasons behind this decline in investment, the

macroeconomic situation is all the more worrying in that our economies have plenty of sectors requiring investment (such as ecological transition, elderly care, education). Investment is spending made today so as to prepare the future. With low levels of investment, our society is no longer preparing its future. With firms reluctant to invest, it may be a great opportunity for the state to assume its role in socializing investment in these sectors, thereby realizing what Keynes imagined in the 1930s.

REFERENCES

Aftalion, A. (1908), 'La réalité des surproductions générales: essai d'une théorie des crises générales et périodiques', *Revue d'économie politique*, **22** (10), 696–706.

Baddeley, M. (2002), 'Investment: accelerator theory of', in B. Snowdon and H. Vane (eds), *An Encyclopedia of Macroeconomics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 369–72.

Bernanke, B., M. Gertler and S. Gilchrist (1999), 'The financial accelerator in a quantitative business cycle framework', in J.B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics*, Vol. 1, Amsterdam: Elsevier Science, pp. 1341–93.

Bhaduri, A. and S. Marglin (1990), 'Unemployment and the real wage: the economic basis for contesting political ideologies', *Cambridge Journal of Economics*, **14** (4), 375–93.

Clark, J.M. (1917), 'Business acceleration and the law of demand: a technical factor in economic cycles', *Journal of Political Economy*, **25** (1), 217–35.

Crotty, J. (2003), 'Core industries, coercive competition and the structural contradictions of global neoliberalism', in N. Phelps and P. Raines (eds), *The New Competition for Inward Investment: Companies, Institutions and Territorial Development*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 9–38.

Dallery, T. (2009), 'Post-Keynesian theories of the firm under financialization', *Review of Radical Political Economics*, **41** (4), 492–515.

Dixit, A.K. and R.S. Pindyck (1994), *Investment Under Uncertainty*, Princeton, NJ: Princeton University Press.

Eichner, A. (1976), *The Megacorp and Oligopoly: Micro Foundations of Macro Dynamics*, Cambridge, UK: Cambridge University Press.

Fisher, I. (1930), *The Theory of Interest*, New York: Macmillan.

Galbraith, J.K. (1967), *The New Industrial State*, Boston, MA: Houghton Mifflin.

Goodwin, R. (1967), 'A growth cycle', in C.H. Feinstein (ed.), *Socialism, Capitalism and Economic Growth*, Cambridge, UK: Cambridge University Press, pp. 165–70.

Gordon, R.J. (2012), 'Is US economic growth over? Faltering innovation confronts the six headwinds', National Bureau of Economic Research Working Paper, No. 18315.

Harcourt, G.C. and P. Kenyon (1976), 'Pricing and the investment decision', *Kyklos*, **29** (3), 449–77.

Harrod, R.F. (1933), *International Economics*, London: Nisbet.

Harrod, R.F. (1939), 'An essay in dynamic theory', *Economic Journal*, **49** (1), 14–33.

Harrod, R.F. (1948), *Towards a Dynamic Economics*, London: Macmillan.

Harrod, R.F. (1951), *The Life of John Maynard Keynes*, London: Macmillan.

Hayashi, F. (1982), 'Tobin's marginal q and average q: a neoclassical interpretation', *Econometrica*, **50** (1), 213–24.

Jorgenson, D. (1963), 'Capital theory and investment behaviour', *American Economic Review*, **53** (2), 247–59.

Kaldor, N. (1966), *Causes of the Slow Rate of Economic Growth of the UK*, Cambridge, UK: Cambridge University Press.

Kalecki, M. (1937), 'The principle of increasing risk', *Economica*, **4** (16), 441–7.

Kalecki, M. (1943), 'Political aspects of full employment', *Political Quarterly*, **14** (4), 322–30.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Koyck, L. (1954), *Distributed Lags and Investment Analysis*, Amsterdam: North-Holland Publishing Company.

Lavoie, M. (1992), *Foundations of Post-Keynesian Economic Analysis*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Marglin, S.A. and A. Bhaduri (1990), 'Profit squeeze and Keynesian theory', in S. Marglin and J. Schor (eds), *The Golden Age of Capitalism: Reinterpreting the Postwar Experience*, Oxford: Clarendon Press, pp. 153–86.

Minsky, H.P. (1975), *John Maynard Keynes*, New York: Columbia University Press.

Robinson, J.V. (1962), *Essays in the Theory of Economic Growth*, London: Macmillan.

Samuelson, P.A. (1939), 'Interaction between the multiplier analysis and the principle of acceleration', *Review of Economics and Statistics*, **21** (2), 75–8.

Skott, P. (1989), *Conflict and Effective Demand in Economic Growth*, Cambridge, UK: Cambridge University Press.

Summers, L. (2015), 'Demand side secular stagnation', *American Economic Review: Papers and Proceedings*, **105** (5), 60–65.

Thirlwall, A.P. (1979), 'The balance of payments constraint as an explanation of international growth rate differences', *Banca Nazionale del Lavoro Quarterly Review*, **32** (128), 45–53.

Tobin, J. (1969), 'A general equilibrium approach to monetary theory', *Journal of Money, Credit and Banking*, **1** (1), 15–29.

Verdoorn, P.J. (1949), 'Fattori che regolano lo sviluppo della produttività del lavoro', *L'Industria*, **1**, 3–10.

Wood, A. (1975), *A Theory of Profits*, Cambridge, UK: Cambridge University Press.

A PORTRAIT OF HENRY ROY FORBES HARROD (1900–78)

Henry Roy Forbes Harrod was born on 13 February 1900 in Norfolk (UK). At first, he did not study economics but was rather interested in philosophy and history. He was appointed Lecturer in Modern History and Economics at the University of Oxford (Christ Church) in 1922. Since Harrod was not very educated in economic matters, he was released from his first two terms by his college so as to improve his knowledge of economics. He was introduced to John Maynard Keynes in Cambridge (UK), where he attended lessons on money and assisted in seminars. Each week, he was also asked to write essays, which would then be discussed with Keynes. This link between Keynes and Harrod persisted beyond Harrod's residence in Cambridge. Harrod repeatedly published articles in the *Economic Journal* while Keynes was editor, and he even took over from Keynes as editor from 1945 to 1961. It was also Harrod who was charged by Keynes's brother, Geoffrey, to write the official biography of John Maynard (Harrod, 1951). Harrod was also one of the readers of the first draft of *The General Theory*, and he worked with Keynes to the preparation of the Bretton Woods conference in 1944.

Harrod published several books and articles on different aspects of our economies, but he also kept a special attachment to his philosophical education, since he published in 1956 a book on the inductive method.

Harrod was one of the main contributors of the Keynesian revolution in the 1930s, but he also helped to extend Keynesian results. Harrod (1948), for example, is an attempt to find out in a dynamic context what Keynes showed in a static environment within *The General Theory*. Beyond these theoretical debates, Harrod also engaged himself in economic policy debates, especially on monetary issues. From time to time, he navigated between political parties, advising the Labour in the 1930s, being member of the Liberal Party for a long time (and even a defeated candidate in 1945), and finally being close to the Conservatives in the late 1950s. Harrod was knighted in 1959 and died in 1978.

Harrod's research works are mainly concerned with economic growth theory. He may even be considered as one of the founders of this specific field of analysis. If Keynes's *General Theory* highlights the difficulties in reaching full employment, Harrod wanted to extend this result to the long run, while showing how hard it is to remain at full employment. The dynamic analysis of the two effects of investment (the capacity effect and the demand effect) shows that the way to full employment is on a knife-edge: either the economy is initially miraculously on this track, or it will permanently move further and further away from this equilibrium situation. While reading Harrod's essay, Keynes answered that it does not seem to him that the economy was totally unstable as implied by Harrod. Instead of this chaotic path, Keynes believed in a corridor of stability:

> It is not impossible that there may be a range within which instability does in fact prevail. But, if so, it is probably a narrow one, outside of which in either direction our psychological law must unquestionably hold good. Furthermore, it is also evident that the

➡

> multiplier, though exceeding unity, is not, in normal circumstances, enormously large. For, if it were, a given change in the rate of investment would involve a great change (limited only by full or zero employment) in the rate of consumption. (Keynes, 1936, p. 227)

Nowadays, heterodox economists carry on this debate on stability. Especially, there is an ever-lasting opposition between Kaleckian authors, who are proponents of stability (see, for example, Lavoie, 1992), and neo-Harrodian authors, who support instability as a more appropriate property of the theoretical models meant to represent our economies (see, for example, Skott, 1989).

A second field where Harrod's heritage is still very relevant concerns international economics. Before the publication of *The General Theory*, Harrod (1933) suggested the existence of an export multiplier, which plays on the same ground as the investment multiplier: an increase in exports will lead to a bigger increase in activity because of the increase in income distribution by exporting firms, which leads their workers to increase their consumption spending, and so on. The static export multiplier of Harrod has been used to underline the strategic component of the industrial basis for development. A full literature has now flourished to analyse economic growth as constrained by the balance of payments (see, for example, Thirlwall, 1979). In a sense, this literature only extends in a dynamic context the proposition made by Harrod (1933) in a static environment.

Harrod's legacy provides key insights in the understanding of our economies' dynamics. Apart from economic growth instability or external constraints, Harrod's research works continue to be quoted in a wide range of economic analyses (for example, Harrod's neutrality of technical progress).

True or false questions

1. In neoclassical models, production factors are substitutable.
2. According to Keynes's view, variations in interest rates allow for the equalization of investment to saving.
3. The marginal efficiency of capital means the same thing as the marginal productivity of capital.
4. If Tobin's Q is higher than one, this means that it is preferable to buy new productive equipment rather than buying a competitor on the stock market.
5. At the macroeconomic level, it is equivalent to buy a competitor or to invest in new equipment.
6. All stock bought on the stock market brings new financing to firms.
7. Harrod's warranted growth rate is always equal to the sum of labour force growth and technical progress.
8. Minsky's paradox of tranquillity describes how the absence of financial crises entails a reduction in risk perception, which leads economic actors to take more risks, thus leading to a new financial crisis.
9. There is a consensus to assess that financialization has led to an increase in investment.
10. An increase in demand will necessarily lead to an increase in the profit rate.

Multiple choice questions

1. In neoclassical theory, firms invest less in the case of an increase in the interest rate because such an increase:
   a) represents an increase in the financing cost of investment;
   b) yields no better return on financial investment;
   c) reduces future demand;
   d) increases uncertainty.
2. According to Keynes, what do firms know when they decide to invest?
   a) Future durability of new equipment.
   b) Future structure of their costs.
   c) Current demand for their products.
   d) Evolution of interest rates.
3. When is it preferable to buy a competitor rather than buying new equipment?
   a) When investment goods prices are high.
   b) When the stock exchange is high.
   c) When the stock exchange is stable.
   d) When investment goods prices are low.
4. The multiplier value is higher if:
   a) investment is high;
   b) investment is low;
   c) the propensity to consume is high;
   d) the propensity to consume is low.
5. The accelerator principle describes:
   a) how variations in investment generate bigger variations in output;
   b) how variations in output generate bigger variations in investment;
   c) how the level of investment determines the evolution of output;
   d) how the level of demand influences the evolution of investment.
6. An increase in demand may cause an increase in investment if:
   a) firms' productive capacities are not fully used;
   b) the increase in demand is only temporary;
   c) firms are financially constrained;
   d) investment goods' firms do not have fully used productive capacities.
7. For Harrod, if firms operate below their target utilization rate for their productive capacities:
   a) they will invest more than the expected growth in demand;
   b) they will invest as much as the expected growth in demand;
   c) their utilization rate will increase;
   d) their utilization rate will decrease.
8. In Kalecki's profit equation, macroeconomic profits are fuelled by:
   a) public deficits;
   b) trade deficits;
   c) workers' saving;
   d) capitalists' saving.
9. In Robinson's banana diagram, an increase in the propensity to save will lead to:
   a) an increase in the growth rate;
   b) an increase in the profit rate;
   c) a decrease in the growth rate;
   d) a stability of the profit rate.

10. In Wood's diagram, the finance frontier moves down if:
a) the interest rate goes up;
b) investment goes down;
c) dividend payments decrease;
d) new equity issues decrease.

10

Aggregate demand

Jesper Jespersen

OVERVIEW

This chapter explains that:

- aggregate demand comprises of private consumption, private real investment, government expenditures on goods and services, and net exports;
- neoclassical economists consider aggregate demand as rather unimportant; they argue that output (gross domestic product) is determined mainly by the supply of labour and capital; the market system is considered as self-adjusting, which makes them conclude that in the longer run 'the supply of goods and services creates its own demand';
- according to heterodox economists and Keynesian macroeconomic theory, aggregate demand is an important analytical concept; it is the major driving force behind the development of output and employment in the short and longer run; this consideration makes demand management policies instrumental for creating macroeconomic stability and economic growth;
- Richard Kahn was one of the first Keynesian economists who contributed to the theory of aggregate demand; he invented the analytical concept of the 'multiplier' as a short-run dynamic phenomenon.

KEYWORDS

- **Aggregate demand:** It is an important theoretical concept defined as the firms' expected proceeds from (future) sales of goods and services. This concept is used intensively by Keynesian macroeconomists as the main causality of changes in output and employment, which is in clear opposition to the neoclassical proposition that 'supply creates its own demand'.
- **Demand management policies:** They are undertaken by governments (fiscal policy) and central banks (monetary policy) at the recommendation

of Keynesian economists with the aim of making the macroeconomic development more stable and to reduce unemployment rates.

- **Effective demand:** Theoretically defined by Keynes as the intersection between two functions representing aggregate demand (expected proceeds) and aggregate supply (normal costs).
- **Full employment:** The level of employment where the demand and supply (curves) for labour intersect. If some people at that 'equilibrium' real wage choose to stay unemployed, they are in neoclassical theory considered as voluntarily unemployed due to for instance generous social benefits, so-called 'natural unemployment'.
- **Income multiplier:** It is the relation between a change in an exogenous part of aggregate demand and the resulting change in output. The size of the income multiplier is determined by the marginal propensity to consume, the marginal tax rate, social benefits, and marginal propensity to import. The magnitude varies between countries and during the business cycle and in modern welfare states usually lies between 1.5 and 0.7.
- **Private consumption and investment:** These are the dominant demand components of aggregate demand and the main drivers of macroeconomic dynamics making the causality run from firms' expected proceeds, via decisions on production to employment in the short run (business cycles) and the longer run (growth).

Why are these topics important?

Macroeconomic analysis is statistically based on the National Accounting System (NAS). The general principles behind the NAS were established in the 1930s as a result of the Great Depression, which caused a steep fall in production and a rise in unemployment. At the time, governments felt an urgent need for more precise information about the overall status of the general economic situation. As a result, they established a standard for the statistical principles of the NAS, since these were accepted internationally, overseen by the United Nations organization and agreed to be followed by its member states.

The basic principles of the NAS are quite simple and aim at giving an objective picture of economic activity in a country. Registrations are based on statistics covering economic transactions, which have a market price and/or create money income (so-called factor income) and add up to gross domestic product (GDP).

Aggregate demand is defined by the NAS as the sum of several specific economic transactions, which adds up to final demand. The NAS is organized in such a way that factor income (wages and profits), GDP and final demand by

definition are made similar. The focus point in macroeconomic[1] analysis is usually GDP and in heterodox economics also aggregate demand, which are important factors determining employment as well as unemployment.

Aggregate demand consists of a number of subcomponents. They are usually defined and categorized within the NAS by, on the one hand, the characteristics of the produced goods and services and, on the other hand, by the actors (households, firms, government sector, or those actors living abroad), who have undertaken the demand (by ordering and purchasing the produced goods and services) registered within the statistics. Aggregate demand refers to items and activities that cause physical production (real economic activities), which employ labour and involve the use of physical capital. As a consequence of this production, these factors receive (factor) income usually in the form of money wages and profits, which eventually are spent in purchasing goods and services.

Aggregate demand (consisting of newly produced goods and services) is divided, within the NAS, into a number of analytically relevant demand components: private consumption, private real investment, government consumption and real investment, and net exports.

Within the private sector, it is assumed that households, mainly wage earners, decide on consumption (C), which (statistically) consists of items such as food, clothes, entertainment, travels, cars and housing rent. Private houses, although usually bought by households, are considered as an investment, because newly produced houses are (like firms' investment) long-lasting. This was discussed at length in Chapter 8.

Private firms decide on real investment (I), which, as a part of aggregate demand, is defined as only newly produced buildings, infrastructure, vehicles and machinery. Likewise, the government contributes to aggregate demand by public consumption and investment (G) in the form of employing people in the public sector and/or demanding goods and services produced in the private sector. Finally, net exports ($X - M$; exports of goods and services minus imports of goods and services) are the part of aggregate demand that is sold abroad minus goods and services bought in foreign countries.

It has become standard to present aggregate demand (AD) in analytical models as follows:

$$AD \equiv C + I + G + (X - M) \tag{10.1}$$

where C represents private consumption, I private real investment, G government consumption and real investment, X exports of goods and services, and M imports of goods and services; $(X - M)$ is thus net exports (NX).

Further, it is a statistical convention that aggregate demand is made equal to the supply of goods and services, which is the gross domestic product (GDP). The national accounting identity is as follows:

$$AD \equiv GDP \tag{10.2}$$

Domestic production is assumed to be carried out by private and public firms or public institutions. *GDP* is physically produced by employing labour and physical capital (land, buildings and machinery) in the economic system. The value of production, *GDP*, measured at market prices, is the remuneration that the factors of production obtain in the form of wages (labour) and profits (physical capital). The sum of wages (W) and profits (P) is called gross factor income (Y).

These statistical conventions bring us to another important national accounting identity, as follows:

$$AD \equiv C + I + G + NX \equiv GDP \equiv W + P \equiv Y \tag{10.3}$$

Identity (10.3) shows that income and output are identically equivalent at the macroeconomic level.

From statistical concepts to macroeconomic theory

Macroeconomic data are important. They allow us to understand the overall economic development during a given period of time. Although the main concern is usually directed towards GDP, the statistical records of the NAS contain very detailed information about many other important macroeconomic variables, such as employment, balance of payments, inflation and public sector budgets, just to mention those variables that usually cause political concern.

The overlying task for macroeconomists is to establish plausible explanations of the statistical development of the major macroeconomic variables through time. For that purpose macroeconomists set up hypotheses about the causalities and dynamic structures within the economic system.

As regards macroeconomic theory, there are two dominant and competing hypotheses of how to explain the evolution of GDP, employment, inflation, and so on. Both theoretical explanations take departure from equation (10.3). In the previous section, this equation was presented as a statistical identity, which is useful in securing consistency within the NAS; but it does not give any information about causality. This is where Keynesian macroeconomists deviate from the conventional neoclassical explanation of how GDP is determined.

To put it briefly, as explained in Chapter 3, neoclassical economists argue that the supply of production factors (labour and capital) together with technological capability (productivity) determine (mainly) the potential level of output and also, under ideal market conditions, the actual level of production. This amounts to saying that neoclassical economists assume that 'the supply of goods creates their own demand' (Say's Law).

By contrast, Keynesian economists argue that it is the aggregate demand for goods and services that makes firms undertake production and employ labour (and capital). In other words, Keynesian economists argue that aggregate demand is a necessary (but not always sufficient) condition to make firms produce (Box 10.1).

Hence, aggregate demand plays a much more prominent role in Keynesian macroeconomic theory than in neoclassical economics. Therefore, the next section, on the role of aggregate demand in neoclassical theory, will be rather brief compared to the discussion of the determinants of aggregate demand in Keynesian macroeconomics, which takes a lot of inspiration from John Maynard Keynes's original contribution in *The General Theory of Employment, Interest and Money* (1936).

Box 10.1

TWO ALTERNATIVE MACROECONOMIC THEORIES (CAUSAL RELATIONS) OF PRODUCTION, EMPLOYMENT AND AGGREGATE DEMAND

1. Neoclassical theory: supply of labour ➔ (full) employment ➔ GDP ➔ aggregate demand.
2. Keynesian theory: aggregate demand ➔ GDP ➔ employment (unemployment).

The neoclassical theory of aggregate demand

Aggregate demand plays an unobtrusive role in neoclassical macroeconomics. As shown in Box 10.1, the assumed causality within the economic system runs from the supply of factors of production via a well-functioning labour market with flexible money wages to the size of GDP. Further, it is assumed within this macroeconomic framework that the supply of goods creates their own demand, which means that aggregate demand for goods and services always adjusts to potential output. Hence, it is not total output, but only the subdivision of output between consumption and investment goods, which is determined by consumers (households), investors (firms), and the general government sector.

The theoretical arguments build on microeconomic principles of individual optimization (utility and profit maximizing behaviour) and single market adjustment (usually under the condition of perfect competition) put into a macroeconomic framework of general (that is, total) market clearing system.

The causal starting point of the determination of aggregate demand in neoclassical economics is the neoclassical (perfectly competitive with a flexible wage level) labour market, where the demand and supply of labour are derived from aggregated individual optimization behaviour (Figure 10.1). Employment is determined by the intersection point of L_d and L_s functions, which stand for the demand and supply of labour respectively. Their intersection marks the equilibrium within the labour market.

Figure 10.1 The neoclassical labour market

Yet, the occurrence of this equilibrium can be impeded by rigidities of real-wage adjustment, caused by external market forces such as a minimum wage set by the government, or market power exercised by trade unions. If the real wage is at a too-high level – that is, above the equilibrium real wage – there will be a discrepancy between labour demand and labour supply: workers will offer more labour than what firms are ready to demand at that real wage, resulting in so-called involuntary unemployment; involuntary because individuals do not have the power of reducing the real wage level (set by external market forces) or the possibilities to improve their employment situation in another way.

In some versions of neoclassical theory, temporary unemployment is also explained by a sluggish adjustment of the real wage owing to a lack of information, employment contracts and social security systems. In other words, the real wage will move toward the equilibrium only slowly, because of frictions or imperfections in the labour market. This adjustment process might even take several years to complete. But the labour market is assumed to eventually reach the market clearing equilibrium, with a natural rate of employment, in the longer run. Output will grow as a consequence of an increased labour supply, because it is assumed in neoclassical theory that supply creates its own demand (Box 10.2).

When the labour market has adjusted to its equilibrium, any further expansion in output is explained within the neoclassical growth model (see Chapter 13). In that model, it is the developments in productivity and labour supply that determine the long-run growth rate of output. Households decide on how to divide factor income between current consumption and savings (deferred consumption). Financial savings are assumed to be transformed

Box 10.2

NEOCLASSICAL MACROECONOMIC THEORY (CAUSAL RELATIONS) OF PRODUCTION, EMPLOYMENT AND AGGREGATE DEMAND

1. Labour market is assumed to adjust to (full) employment due to a flexible real wage.
2. (Full) employment determines the level of output (GDP).
3. Supply of output (GDP) creates its own (aggregate) demand for goods and services.
4. Behaviour of households and firms determines the subdivision of aggregate demand into consumption and real investment.

automatically into productive real investment via a perfectly functioning credit market. In this model, savings are always identical to investment, and therefore do not cause any drain on aggregate demand. Further, it is assumed that savings determine investment.

Neoclassical aggregate supply: the aggregate supply (AS) model

Neoclassical models of output determination have their causal starting point on the labour market. As discussed above, the crucial assumption is that, owing to market forces (perfect competition), the labour market will adjust by itself to an equilibrium, that is, the intersection point of L_d and L_s curves in Figure 10.1, the point which Milton Friedman dubbed the 'natural rate of unemployment' (that is, no involuntary unemployment). Given the labour market structural conditions and the social security system, unemployment cannot be reduced further without causing wage and price inflation. Hence, the longer-run output level (the long-run aggregate supply, LRAS, curve in Figure 10.2) is determined by labour market equilibrium and is assumed to be constant (while disregarding changes in capital stock and productivity).

If firms erroneously expect the price level of output in the next period to increase, they will react as if the real-wage level (cost of production) has fallen, since an increase in the price level decreases the real wage and hence increases the profit rate. As a result, they will employ more people and expand output according to their short-run aggregate supply (SRAS) curve (in Figure 10.2). Employment is, in this case, increased temporarily and unemployment reduced below its natural level, which will shortly make money wage begin to rise, thereby increasing the real wage. Firms will realize at the end of the period that their expectations had been wrong, since the

Figure 10.2 The neoclassical output model

real wage has not fallen as much as they expected. They will start therefore to reduce output and employment in the following period in accordance with the new and correct information on prices and wages. At the end of the adjustment process, output and employment will be unchanged (predetermined by the long-run aggregate supply, LRAS curve); but the price level might have increased permanently.

Something similar is assumed to happen if governments try to increase the level of employment by expansionary monetary or fiscal policies (see Chapters 7 and 12). Increased public sector employment might temporarily increase total output; but money (and real) wages go up and squeeze private sector profit and production, leaving the overall GDP (private and public output) unchanged at the end of the adjustment period. The same labour market development occurs when increased public investment is undertaken. Public investment drains private savings and raises the rate of interest (see Chapter 12). Hence, private investment will fall consequently. In neoclassical economics this mechanism of public expenditures substituting private consumption or investment is called the crowding-out effect. It changes the composition of aggregate demand, but leaves total GDP (output) unchanged.

In general, any beneficial effects of expansionary demand-management policies are disputed by neoclassical economists. They claim that the market system in principle is self-adjusting. Hence, in the case of unemployment there is no need to undertake an expansionary demand policy. Rather, the economic system and especially the labour market should be made more flexible, and the wage level more responsive to unemployment. Neoclassical economists claim that an output gap is not caused by a lack of effective demand but is more often caused by rigidities in the labour market, which therefore should be reformed.

The Keynesian theory of aggregate demand and output

Changes in output are decided by firms (and governments) in light of effective demand, which means the expected sales of goods and services by private firms produced at a positive profit margin. Expectations concerning the future are always based on uncertain information. Proceeds from future sales cannot be calculated or known with certainty. In practice, these expectations are subjectively formed partly on historical data and partly on vague information on what might happen in the future, often influenced by waves of optimism and pessimism.[2] On the other hand, a number of firms produce only in response to specific orders from other firms or from consumers or the

Box 10.3

SUBDIVISION OF AGGREGATE DEMAND FROM A KEYNESIAN POINT OF VIEW

$$AD = C + I + G + (X - M) \qquad (10.1)$$

1. Private consumption, C, is determined by the marginal propensity to consume out of disposable income, Y_d.
2. Private savings, S, is the residual part of disposable income not spent on consumption.
3. $C + S \equiv Y_d$.
4. Private (real) investment, I, undertaken by firms is based on rather uncertain expectations about future demand and proceeds.
5. Government expenditures on goods and services, G, are often determined by economic policy, so-called demand management policy.
6. Foreign trade (balance of payments) consists of export of goods and services minus import of goods and services, $(X - M)$, and is determined by preferences, relative growth rates and international competitiveness.

public sector. In these cases it is the actual demand that gives rise directly to production and employment (Jespersen, 2009).

The statistical aggregation of these millions and millions of daily and monthly transactions caused by households' consumption and firms' investment (domestically and internationally) will, together with information about likely future events, comprise the backbone of expected aggregate demand and will, therefore, dominate the firms' production decisions. History and qualified information are essential for the formulation of the aggregate demand function in a Keynesian macroeconomic model (Box 10.3).

Private consumption

Private consumption is the largest of the subcomponents of aggregate demand, accounting for more than half of GDP. Disposable current household income (factor income plus social benefits minus taxes) is assumed to be the major determinant of private consumption. The population is on average assumed to spend a certain and relatively stable proportion of its current income on consumption. The propensity to consume determines the division of household income into current consumption and savings in the aggregate consumption function:

$$C = C_0 + c_1 Y_d \qquad (10.4)$$

where c_1 represents the marginal propensity to consume,[3] Y_d is disposable income (factor income (Y) + social benefits – taxes), and C_0 is that part of consumption which does not depend on income, which is referred in most textbooks as autonomous consumption.

Of course, private consumption is influenced by many factors other than just disposable income. It has become common practice in recent years also to add a wealth variable, *WE*. In other words, consumption today is assumed to depend on current disposable income and wealth. The wealthier society becomes, the larger a share of current income it can afford to spend:

$$C = C_0 + c_1 Y_d + c_2 WE \tag{10.5}$$

Further, distribution of income, as emphasized by the Polish economist Michał Kalecki, and credit facilities also matter with regard to private consumption. The variable C_0 in equation (10.5) represents all these factors besides disposable income and wealth, which make the consumption function shift from time to time. These factors are many and partly unknown, partly unstable and partly of minor importance.

The part of private disposable income that is not used for consumption is saved. Hence, savings are income not spent on consumption, and instead put into banks as deposits or into pension schemes for consumption in old age. The point here is that savings do not form a part of aggregate demand. In fact, the immediate effect of increased financial savings is a reduction of private consumption (an important part of aggregate demand), which has a negative impact on output. Hence, the larger the propensity to save, the smaller the propensity to consume, because private consumption (C) and private savings (S) add by definition to disposable income. In other words, the marginal propensity to consume and the marginal propensity to save must add to 1. Hence, we have:

$$C + S \equiv Y_d \tag{10.6}$$

Private investment

Private firms undertake investment when they buy new machinery or vehicles, construct new buildings for factories or offices, or create infrastructure. In macroeconomics, the word 'investment' is reserved for economic activities that increase the stock of physical capital. This is true for three analytical reasons:

1. Investment gives rise to output (and employment) and is part of aggregate demand. Hence, investments are recorded within the NAS as newly physically produced goods that are a part of GDP and therefore create factor income. In this respect, investments in the NAS are very different from financial investments (for instance, when one buys stocks on the stock exchange) or the purchase of already-produced physical capital, not to mention land (and other non-producible assets).
2. Private investments and private savings are two absolutely distinct economic activities, undertaken by two different groups in society. Firms undertake investment, and do so to be more productive; whereas saving is done (mainly) by households who save for old age or for the uncertain future. Investment is part of aggregate demand; savings is non-consumption and therefore not part of aggregate demand. This may lead to what is called the 'paradox of thrift': increases in the propensity to save out of disposable income by private households (and firms) lead to a decrease in private consumption, aggregate demand and, hence, output.
3. Since investment gives rise to newly produced real assets, once it has been undertaken, the size of physical capital will increase. As such, the productive capacity of the economic system increases, which in turn increases the growth potential of GDP (see Chapter 13).

As stated above, private firms undertake investment, and do so with the purpose of making a profit in the future. Hence, they calculate what they can expect to add to their profits in the future by undertaking investments in physical capital today. These calculations, however, are very uncertain. Over and over again during the chapters in *The General Theory*, Keynes emphasized that such calculations are by nature uncertain and will at the end depend quite a lot on the psychology of the manager and of the business mood in general. Bursts of optimism and pessimism spread in society with a self-reinforcing effect owing to, among other things, what Keynes called 'animal spirits'.

So why would a firm undertake investment? New capital is needed to substitute used and worn-out capital goods. Real assets do not last forever, and new technology requires that the existing capital equipment is renewed and updated. These investments replace old capital and are called amortization; they represent the largest share of private investment in the NAS. Yet, this has no impact on the size of the physical capital stock: amortization simply replaces old machines and equipment.

Although amortization goods (reinvestment) are produced and add to GDP, they are, as mentioned, only replacements for worn-out capital. If

amortization investment did not take place, the capital stock would shrink and society would eat up its physical wealth. A number of amortization decisions are made rather mechanically, because a broken machine, an outdated vehicle or an old-fashioned technology has to be replaced to make the economic system productive and competitive. Because of this, amortization is a part of the costs of production during the period under consideration, which firms have to pay. Therefore, it has become practice to separate gross profit and net profit in the NAS, where net profit is gross profit minus amortization.

By contrast, positive net investments enlarge the size of the real capital stock. As explained above, they rely more heavily on uncertain expectations of future proceeds from an expanding aggregate demand. The foundation of this calculation regarding the future is by nature quite uncertain and depends on the state of confidence: 'There is, however, not much to be said about the state of confidence a priori. Our conclusions must mainly depend upon the actual observation of markets and business psychology' (Keynes, 1936, p. 149).

Real investments are long-lasting and can usually not be financed only by current profits. Therefore, when firms have made up their mind with regard to the expected return on new investments, they have to confront the expected return with the rate of interest, which they have to pay to obtain the necessary finance. This suggests two things: first, investment depends on the difference between the expected rate of return and the banks' rate of interest; second, monetary policy as practiced by central banks could have an influence over investments by making credit facilities more easily available and by reducing the rate of interest (see below, and Chapter 7).

Government expenditures on goods and services

As we can see from equation (10.3), government expenditures, G – that is, when governments consume (goods and services) and invest – are part of aggregate demand. In the 1930s their share was less than 10 per cent of total aggregate demand. Today, depending on the country, this share has risen to 20–30 per cent. Unfortunately, macroeconomic theory does not have that much to say about this expansion of the public sector, because it was mainly initiated by political considerations, especially when building up the welfare state. The government demand for goods and services when demand management policies are analysed will be considered below, and in Chapter 12.

Another important part of public expenditures is social income transfers (benefits and subsidies), which do not appear directly in the definition of

aggregate demand. This is so because these expenditures are not paid to households in exchange for a productive activity. They are instead paid according to specific social criteria; of course, usually with the aim of protecting people against material starvation. In macroeconomic analysis, social benefit is considered a part of disposable income, Y_d, in equations (10.4) and (10.5), and will therefore have only an indirect impact on aggregate demand, when these social transfers are spent on private consumption.

Public expenditures are mainly financed by taxes (income, consumer and real-estate taxes), which, similarly to social income transfers, do not appear in the definition of aggregate demand, because a tax is by definition not a payment for a specific productive activity. On the other hand, taxes reduce disposable household income and thereby reduce private consumption. Although the impact on aggregate demand and output is indirect, a change in the tax rate is considered as one of the major instruments of demand management policies, as explained below.

International trade

Looking at equation (10.3) two items related to international trade appear: exports and imports of goods and services. They both have a direct impact on GDP. In larger countries such as the United States, exports amount to approximately 20 per cent of GDP; whereas in small open economies such as the Netherlands, exports come close to 80 per cent of GDP.

Exports are determined by the demand for goods and services by households and firms in other countries. For instance, when Chinese importers purchase goods produced in the United States (US), these are considered exported goods in an equation determining US aggregate demand. Increased purchasing power abroad, especially in neighbouring countries, has a positive impact on exports: when households in other countries consume more of our goods, we export more goods. Similarly, domestic aggregate demand is not only directed towards domestically produced goods and services: we also consume goods and serviced produced abroad; imports therefore play a significant role and thereby reduce the total aggregate demand.

Some goods (such as coffee and bananas) cannot be produced in all countries owing to climatic conditions. Some countries are too small to have efficient production in all industries (cars, aeroplanes, advanced electronics, and so on). But even in countries that have their own production of nearly all goods, one may see apparently similar goods being imported and exported. Wine from all over the world can be bought also in wine-producing countries.

Cars are imported and exported in and out of all major industrial countries. The reason for this is partly specific consumer preferences, partly different costs of production and of delivering. Depending on the type of product (or service) under consideration, its costs of production have a rather important impact on firms' international competitiveness, which especially in the longer run will make an impact on trade flows in and out of the country, and therefore on its output. Imbalances between exports and imports show up in the current account of the balance of payments. A surplus has a positive impact on GDP (and employment), because exports exceed imports in that case.

The income multiplier

The original simple multiplier

The marginal propensity to consume (see equations 10.4 and 10.5) plays a major role in all macroeconomic models. It determines the expansionary effect of increased income inside the system. When an additional unit of income is earned, the propensity to consume tells us the share of this additional income that is spent on consumption. The additional consumption has to be produced, which we know from equation (10.3) creates an equivalent amount of factor income. A part of this additional income is spent on consumption creating more factor income, which is partly spent, and so on. This expansionary effect, however, will not go on forever. Each time income is created, a part of it is saved and therefore withdrawn from aggregate demand. This dynamic development is called the 'income multiplier'. It was first introduced into macroeconomic theory by Richard Kahn, and is illustrated in Box 10.4.

The calculation of the income multiplier within a simple macroeconomic model, where only private consumption is considered, is quite simple. The multiplier effect depends only on the propensity to consume. The larger the propensity to consume, the more income is spent on consumption (and

Box 10.4

THE SIMPLE INCOME MULTIPLIER

$$\text{MULT} = 1/(1 - c) \text{ or } \text{MULT} = 1/s \qquad (10.7)$$

where c is the marginal propensity to consume out of disposable income and s the marginal propensity to save out of disposable income, because $c + s \equiv 1$.

production) each time an additional unit of income is earned, hence the larger the multiplier becomes.

For instance, if the marginal propensity to consume is estimated to 0.9, then the income multiplier will be equal to 10. This means that each time a household receives an additional $100, it will spend an $90 on consumer goods and services. When these $90 are spent on output, they become someone's else income, who is assumed to spend 0.9 x $90 = $81 on consumption. These $81 become someone's else income, who will spend 90 per cent of that, hence $72.90, and so on. As one can see, an initial increase in consumption, which increases output, has a 'multiplying' effect on output. The theoretical explanation of the multiplier principle is one of Keynes's most fundamental contributions to macroeconomic analysis.

The elaborated multiplier

Modern welfare states are primarily financed by income and consumption taxes. Since taxes are paid out of current income, tax revenues (net of social benefits) reduce disposable income and thereby private households' consumption, which is thus a drain on aggregate demand. The higher the income (and consumption) tax rate, the larger is the drain. In this respect taxes, like private savings, will reduce the size of the multiplier, because an additional unit of factor income is taxed quite highly, in some countries by 40 or 50 per cent. This tendency of a diminished income multiplier is reinforced further by social benefits being dependent on changes in unemployment, which increases the marginal net tax rate, *nt* (the tax rate plus the unemployment benefit rate). Hence, within modern welfare states the size of the income multiplier is reduced substantially compared to the 1930s, and much more in the Scandinavian countries than in the United States.

Further, the income multiplier is reduced by imports of goods and services. When domestic aggregate demand is increased, especially private consumption and firms' investment, imports will go up and some income will thus leak abroad and cause no further domestic multiplier effect. Hence, the higher the marginal propensity to import, *m*, the smaller the income multiplier becomes (Box 10.5).

In general, small countries are more dependent than big countries on imports, owing to both international specialization and economies of scale. This means that the marginal propensity to import is usually larger in smaller countries than in larger and more self-sufficient countries. The higher the marginal propensity to import, the smaller the income multiplier becomes. In a number of the smaller countries with rather extended welfare states, the

Box 10.5

THE ELABORATED, BUT STILL SIMPLE, INCOME MULTIPLIER

$$MULT = 1/(1 - c + nt + m) \text{ or}$$
$$MULT = 1/s + nt + m \qquad (10.7)$$

where:
c is the marginal propensity to consume out of disposable income;
s the marginal propensity to save out of disposable income, because $c + s \equiv 1$;
nt is the marginal net tax rate (taxes + subsidies);
m is the marginal propensity to import.

A numerical example from a modern welfare state, such as the Netherlands, is as follows:

$$c = 0.9$$
$$nt = 0.5$$
$$m = 0.3$$
$$MULT = 1 / (0.1+0.5+0.3) = 1 / 0.9 = 1.1$$

This simple example shows that the income multiplier has been reduced considerably due to the development of the modern welfare state and increased international trade.

This is a double-edged development: (1) when an economic crisis erupts, the modern welfare state is much more robust than in the 1930s; but (2) if unemployment has gone up, it requires a much larger economic policy impact to reduce the rate of unemployment, as in Greece since 2009.

income multiplier has become smaller than one, owing to high taxes, social benefits and imports.

Expected aggregate demand and supply: effective demand

Aggregate demand (*AD*) represents the proceeds producers as a whole expect from future sales. The slope of the *AD* curve in Figure 10.3 is determined by the marginal propensity to consume out of disposable income (see equations 10.4 and 10.5), where disposable income is defined as factor income (*Y*) plus social benefits minus taxes, as explained above.

The aggregate supply (*AS*) curve represents the variable costs of production (wages and raw materials) plus a normal business profit.[4] The cost-determined *AS* curve bends upwards, following the convention that in the short run it becomes more expensive to increase output, because firms have to pay extra wages when labourers work more than normal hours,[5] and sometimes less-efficient machines are used for (quickly) increased production. In the longer run the shape of the cost curve might be quite different, depending on new technology and increased productivity (see Chapter 13).

Figure 10.3 The principle of effective demand (short run)

The intersection point between the *AD* curve and the *AS* curve represents the short-run equilibrium level of output, which fulfils the producers' expected proceeds and gives a normal rate of profit. Keynes called this intersection point 'effective demand'. When producers have decided on output for the coming period, they also know how many people to employ. Hence, employment is determined by effective demand, which is represented in Figure 10.3 by the vertical line connecting output of goods and services (upper part of the figure) with the labour market (lower part).

In the lower part of Figure 10.3, illustrating the labour market, the conventional short-run demand (L_d) function for labour has also been drawn. Keynes accepted in many ways the relevance of this very conventional

labour-demand curve, assuming individually profit-maximizing firms, decreasing productivity of labour and perfect competition among producers in the labour market. Under these admittedly very unrealistic assumptions the real wage would be determined by effective demand for output following the arrows marked '*A*'. When producers and/or trade unions execute market power in the wage-setting process, which they often do, the real wage is not only determined by market forces, but also dependent on the relative strength of labour market organizations and of the size of unemployment (see Chapter 11).

An expansionary demand management policy causes an upward shift in the *AD* curve in Figure 10.3. The new intersection point between aggregate demand and aggregate supply determines how much output and employment will increase according to the model. It is the slope of the *AD* curve that determines the dynamics of the adjustment process; that is, the size of the income multiplier. The steeper the *AD* curve, the larger the multiplier.

Demand management policies

As discussed earlier, post-Keynesian economists do not share the neoclassical precondition of a self-adjusting market mechanism that moves the economic system toward full employment. Without this mechanism, the economy can easily get stuck at a high level of unemployment. This was precisely the case in the interwar period, when unemployment remained high for long periods of time. This cannot be justified using neoclassical economics, which sees unemployment as only transitory. Keynes, however, saw neoclassical theory as a special and usually irrelevant case, because the general case was persistently high unemployment. This explains the historical context in which Keynes wrote his *General Theory*.

Yet, empirical evidence shows that periods of persistent unemployment are not limited to the interwar period. In fact, periods with a high rate of unemployment have occurred since the early 1970s both in Europe and in the United States, although somewhat more fluctuating in the latter. Hence, market-based economic systems do not converge to a low rate of unemployment by themselves.

In these long periods when the rate of unemployment is above any reasonably defined level of full employment, there is a gap between actual and potential output. Private effective demand is not enough to fill that gap. In these long-lasting periods, demand management policies can be useful to expand aggregate demand by monetary and/or fiscal policies.

Monetary policy is conducted by the central bank. The main monetary policy instruments are the short-term rate of interest, extended credit facilities, and open-market operations (buying long-term bonds from banks, sometimes called quantitative easing, QE) with the aim of reducing the long-term rate of interest (see Chapter 7). Lower rates of interest and easier access to credit are expected to have a stimulating effect on private investment, and to a lesser extent on private consumption. Further, countries with a flexible exchange rate regime may experience a currency depreciation as a consequence of lower rates of interest and QE. In this case, exports go up and imports go down, which will increase aggregate demand and domestic output. A lower rate of interest will also reduce public sector interest payment, which will relieve liquidity-constrained public sector budgets.

Fiscal policy (public expenditures and taxation) is undertaken by governments to correct macroeconomic imbalances in the private sector, especially when unemployment is high. In such cases the government can increase public expenditures (G in equation 10.4) and/or lower taxes and/or increase social spending. These kinds of expansionary fiscal policy will have a positive, but quite different impact on aggregate demand, depending on what instruments are used. There is a vast difference in the size of the income multiplier related to change in specific policy instruments. Usually, public spending for consumption and investment is considered as the instrument with the largest multiplier effect, because the first-round income effect is mainly domestic with very little content of imports. Similarly, social benefits have a rather high multiplier, because very little, if anything at all, is saved in the first round. People with low income spend what they earn, whereas tax reliefs to people with high income have little impact on consumption, output and employment.

Unfortunately, demand management policies are not always simple to undertake. There is a time lag from the adoption of the relevant policy instrument until private firms revise their production plans. In general, monetary policy can be decided relatively quickly, contrary to discretionary fiscal policy, which has to be approved by parliament. In addition, overambitious full-employment policy – such as that undertaken in the late 1960s in Western countries – could, as Kalecki (1943) explained, lead to a change in the working of the labour market and cause rising inflation. Furthermore, expansionary fiscal policy will often leave the public budget with a deficit, which has to be financed by increasing the money supply and/or selling bonds to the private sector (or abroad) (see Chapter 12).

The importance of aggregate demand in macroeconomic analysis was one of the theoretical novelties exposed by Keynes's *General Theory*. It was a long-

lasting process, because in the late 1920s Richard Kahn had already demonstrated, by using public investments (in housing) as an example of active demand management policies, how the government could counterbalance and correct an imbalance in the private sector between savings and private investments, if the rate of unemployment was unacceptably high (see Kahn, 1984).

NOTES

1 In fact, the expression 'macroeconomics' related to the aggregated national accounting items was not established until the second half of the 1930s.
2 'State of confidence' is an expression often used by Keynes in this regard.
3 It is called 'marginal' because it settles how much consumption will increase, if disposable income goes up by one extra (marginal) unit of income.
4 'Normal business profit' is not a well-defined concept, but is related to the ruling rate of interest plus a risk premium, which might vary from industry to industry.
5 A more thorough argument at the macroeconomic level would refer to the Phillips curve. Here it is assumed that the lower the rate of unemployment is, the higher is the rate of wage inflation causing the cost level to rise (see Chapter 11).

REFERENCES

Jespersen, J. (2009), *Macroeconomic Methodology: A Post-Keynesian Perspective*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Kahn, R.F. (1929), *The Economics of the Short Period*, Cambridge, UK: King's College.

Kahn, R.F. (1931), 'The relation of home investment to unemployment', *Economic Journal*, **41** (162), 173–98.

Kahn, R.F. (1984), *The Making of Keynes's General Theory*, Cambridge, UK: Cambridge University Press.

Kalecki, M. (1943), *Studies in Economic Dynamics*, London: Routledge, 2016.

Keynes, J.M. (1930a), *A Treatise on Money, Volume I: The Pure Theory of Money*, London: Macmillan.

Keynes, J.M. (1930b), *A Treatise on Money, Volume II: The Applied Theory of Money*, London: Macmillan.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

A PORTRAIT OF RICHARD KAHN (1905–89)

Richard Kahn was educated in mathematics, physics and economics at the University of Cambridge, United Kingdom. He was associated with King's College from the very beginning. He was elected Fellow in 1930 and succeeded Keynes as Bursar in 1946. He had appointments in the Faculty of Economics and Politics from 1933, and became a full Professor in 1951, at which point he succeeded Dennis Holmes Robertson.

Kahn was Keynes's closest collaborator while Keynes wrote *The General Theory of Employment, Interest and Money*. Kahn was one of the five members of the so-called Cambridge Circus, together with Joan Robinson, Austin Robinson, Piero Sraffa and James Meade, whose principal mission, starting in the autumn of 1930, was to discuss Keynes's newly published *A Treatise on Money*, a two-volume book (Keynes, 1930a, 1930b). Kahn reported once a week to Keynes on the findings of the 'Circus'.

Kahn's most notable contribution to economic analysis was his principle of the income multiplier. This multiplier is the relation between an increase in aggregate demand and the related increase in output. His findings on the multiplier were first published in his seminal 1931 article, 'The relation of home investment to unemployment' in the *Economic Journal*. This paper was developed from his dissertation on *The Economics of the Short Period*, submitted to King's College in 1929. Keynes quickly picked up the idea of the multiplier and elaborated on it in several cases before he made it an integrated element of his *General Theory.*

During the Second World War, Kahn worked as a civil servant in a number of different ministries, which gave him a lot of practical experience concerning real-world economics, planning and the role of economic policies. In 1946 he returned to Cambridge. Although he was appointed full Professor in 1951, he did not really make a lasting impact on the Department of Economics and Politics. When he retired in 1972 his Chair in Economics was taken over by the neoclassical economist Frank Hahn.

? EXAM QUESTIONS

True or false questions

1. Aggregate demand is determined by the supply of labour and real capital.
2. Aggregate demand consists of private consumption, real (private and public) investment, public consumption and net exports.
3. Keynesian macroeconomists assume that GDP is determined by aggregate demand.
4. Neoclassical macroeconomists assume that 'supply of labour creates its own demand'.
5. Disposable income earned by households is the sum of private factor income and social transfers.
6. The propensity to consume out of disposable private income is usually assumed to be larger than one.
7. Improved technology – that is, increased labour productivity – increases the number of people employed.

8. Unemployment is caused by a too-small aggregate demand.
9. The size of the income multiplier depends on the income tax rate.
10. 'Private investment' entered into the national accounting system consists of land, new machinery and buildings.

Multiple choice questions

1. Heterodox economists consider aggregate demand as most important for:
   a) potential GDP;
   b) actual GDP;
   c) labour supply;
   d) the balance of payments.
2. Why do macroeconomists disagree?
   a) Because macroeconomics is not an exact science.
   b) Because macroeconomic explanations differ depending on ideological views.
   c) Because macroeconomists have different views on the role and potential impact of demand management policies.
   d) All of the above.
3. What does the national accounting identity tell us?
   a) GDP is the main goal of economic activity.
   b) Aggregate demand is equal to GDP.
   c) The definition of full employment.
   d) How economic growth is measured.
4. What is the main determinant of private consumption?
   a) GDP.
   b) Aggregate demand.
   c) Disposable income.
   d) Economic policy.
5. What is the simple income multiplier?
   a) The relation between GDP and employment.
   b) The relation between a change in an exogenous demand variable and GDP.
   c) The relation between factor income and disposable income.
   d) The relation between disposable income and the income tax rate.
6. What are real private investments?
   a) The number of government bonds bought by households.
   b) Public infrastructures built by private enterprises.
   c) Buying private land for city development.
   d) Building private houses and factories.
7. How can endogenous business cycles be explained?
   a) By economic policies.
   b) By coronavirus pandemic and other shocks.
   c) By changes in private investments.
   d) By Chinese exports.
8. What is the main determinant of total employment?
   a) The wage level.
   b) Unemployment.
   c) Supply of labour.
   d) GDP and productivity.

9. What causes unemployment?
a) Too little aggregate demand.
b) Too-high wage level.
c) Too many robots.
d) Too many people.

10. The macroeconomic impact of expansionary fiscal policy is:
a) reduced unemployment;
b) increased actual GDP;
c) increased trade deficits;
d) all of the above.

11

Inflation and unemployment

Alvaro Cencini and Sergio Rossi

OVERVIEW

This chapter:

- provides a correct definition of inflation;
- shows that a distinction must be made between inflation and the cost of living;
- shows that the consumer price index is a poor indicator of the rate of inflation and that inflation cannot be identified with a persistent rise of the general price level;
- explains the monetary–structural origin of inflation;
- introduces Keynes's distinction between voluntary and involuntary unemployment;
- shows that involuntary unemployment is pathological and cannot be imputed to economic agents' behaviour;
- proves the existence of a close relationship between involuntary unemployment and deflation, and points to the role played by both the rate of interest and the process of capital accumulation in the rise of unemployment;
- shows that inflation and unemployment are twin effects of the same cause;
- outlines a reform that can solve both pathologies and enable a transition from a disorderly to an orderly system of payments.

KEYWORDS

- **Frictional unemployment:** The unemployment that occurs owing to the normal working of an economic system, where some people are temporarily unemployed until they are able to find another job.
- **Inflation:** Caused by a pathological change altering the relationship between national money and national output, inflation is a disorder causing the loss in the purchasing power of each money unit.
- **Involuntary unemployment:** The pathological unemployment caused by a monetary disorder arising from the actual process of fixed capital formation and the ensuing deflation.
- **Standard of living:** Measures the purchasing power of households over a basket of domestic and foreign commodities. It must be clearly distinguished from inflation.
- **Voluntary unemployment:** The frictional unemployment accompanying the normal working of our monetary economies of production.

Why are these topics important?

The negative consequences of inflation and unemployment should be self-evident. The economic and social problems induced by unemployment affect an increasing number of industrialized countries. The consequences of this disorder are well known: increase in poverty, spending cuts in public services, increased indebtedness (public and private), and social unrest are only some of the negative consequences endured by populations all over the world.

Inflation is only superficially a less disruptive disorder. If we bear in mind that inflation brings about a decrease in the purchasing power of money, we can easily realize that it affects the standard of living of all those people whose income is fixed over time (such as pensioners, beneficiaries of public assistance, owners of non-indexed bonds, and so on).

The failure to address these pathologies correctly has led to a worsening of the disorder, of which the economic and financial crisis that erupted in 2007–08 is a dramatic manifestation at the global level. Inflation and involuntary unemployment are in fact at the core of this macroeconomic disorder, and a correct understanding of their cause is an essential step towards a new, orderly system.

Inflation

The traditional analysis of inflation

Definition

Conventionally, economists define inflation as a persistent increase in the general price level. This means that whenever the price level increases over time, it signals that some inflationary pressures have been pushing a variety of prices upwards. Defining inflation by simply referring to changes in price levels implies that inflation is a result of supply and demand forces: if, on any given market for produced goods or services, there is an increase in demand and/or a reduction in supply, and if the elasticity of supply is low (that is to say, supply does not easily adjust to increases in demand), the prices of the relevant items will be higher as a result of the price determination process. Now, since there are thousands of goods and services, whose prices may move up and down at different rates of change in any given period, inflation is usually measured by referring to an index number, that is, a numerical index that is meant to subsume a great number of items whose prices may change differently during a given period of time (say, between last year and this year).

Measurement

Generally speaking, national statistical offices measure inflation through the Laspeyres price index. The relevant formula is as follows:

$$P_L \equiv \frac{\sum_{i=1}^{n} p_i^1 q_i^0}{\sum_{i=1}^{n} p_i^0 q_i^0} \times 100 \qquad (11.1)$$

where P_L is the Laspeyres price index, i represents a good or service (for $i = 1, \ldots, n$), p_i^1 is the price of item i in the current period (1), p_i^0 is the price of the same item in the reference period (0), and q_i^0 is the quantity of item i purchased in the reference period.

Now, the Laspeyres index 'assumes no consumer substitution occurs in response to changes in relative prices, an assumption that is extreme, unrealistic and unnecessary' (Boskin et al., 1998, p. 7). In fact, when the prices of substitutable goods change over time, consumers tend to substitute the good whose price has increased more with the good whose price has increased

less. Suppose, for instance, that the price of beer X is increasing more than the price of beer Y: assuming that these beers are similar, one is led to reduce or stop purchasing beer X and to increase consumption of beer Y. If so, then the Laspeyres index does not capture this shift in consumers' behaviour, because it considers the quantities of beers X and Y bought in the reference period (0) rather than those purchased in the current period (1). Generalizing this stylized example allows one to understand that the Laspeyres index leads us to overstate the upward movement of prices (see Rossi, 2001/2003, Ch. 1 for analytical elaboration on this).

Causes

There are several causes of an increase in price levels. Neoclassical economists focus largely on the demand side of the economic system, and point out the following:

- Inflation is said to occur whenever the public sector increases government expenditure, as it exerts an upward pressure on the price level by increasing demand on the market for produced goods and services. This implies that, prior to government spending, there was a full employment equilibrium on the goods market: if the public sector is now demanding more goods or services, their market prices will increase, since there can be no general increase in output once all production factors are fully employed.
- Inflation would occur whenever consumers change their forms of behaviour abruptly, and decide to save less in order to consume more goods or services. This may occur, for instance, after a war, when people are relieved to see that the situation has improved and are therefore keener on spending rather than saving. On the assumption that full employment prevails, such an increase in consumption brings about an increase in a number of market prices, hence also an increase in the relevant price level.
- Inflation, it is also claimed, may occur when either the public sector or the private sector obtains more credit from the banking system. In both cases, demand on the goods market increases, leading to an increase in the general price level or, at least, in expected inflation. For instance, the dramatic increase in the money supply as a result of the unprecedented intervention by several central banks in the aftermath of the global financial crisis that erupted in 2007–08 has led many to fear a resurgence of inflation in those countries where monetary authorities have carried out either credit easing or quantitative easing programmes, most notably across the euro area and the United States (US) (see MacLean, 2015; Werner, 2015).

These causes of inflation, in neoclassical economics, are closely linked to the quantity theory of money, which considers that any increase in the money supply leads to some increase in the relevant price level over time (see Mastromatteo and Tedeschi, 2015). In short, orthodox economists believe that the quantity of money and the amount of credit are supply-driven and determined by the central bank. Monetary policy should therefore aim at price stability on the market for produced goods and services, and nothing else.

Now, heterodox economists take a different position. Although many of them adhere to the definition of inflation by its effect on the price level, these economists argue that the causes of inflation are to be found in the never-ending struggle over income distribution between firms (profits) and workers (wages). This is the conflict theory of inflation (see Rowthorn, 1977). As Palley (1996, p. 182) observes, this approach stems from the cost-push theory of inflation developed in the 1950s. In fact, the most widely mentioned source of 'cost-push inflation' arises from 'inconsistent claims on income that emerge from the income distribution struggle between workers and firms' (ibid., p. 182). In this view, the two parties' claims on income may exceed available output in the economy as a whole. So, the excess of income claims over national output originates an increase in prices on the market for produced goods and services. Wage earners try to counteract this upward pressure by bidding for higher wages, thereby setting forth a price–wage or wage–price spiral in which each party seeks to obtain, or to maintain, its targeted income share (see Rossi, 2001/2003, Ch. 5 for analytical elaboration on this).

Some criticisms

Even though the conventional definition of inflation seems uncontroversial, it remains superficial, as it focuses on a surface phenomenon, that is, an increase in some relevant price level. In fact, essentially, inflation is a decrease in the purchasing power of money, as a result of which the general price level increases.

This essential definition needs to be clarified. In particular one has to make clear where the purchasing power of money comes from, and what are those goods and services that determine the purchasing power of any given national currency.

How is it that a national currency can be endowed with a positive purchasing power? The answer is straightforward: it is production that gives a purchasing

power to money. One may indeed have a lot of money, but if there is nothing to purchase (as a result of production), one has in fact no purchasing power at all. This simple thought experiment is enough to establish that the purchasing power of any national currency is determined with regard to the goods and services produced by the corresponding national economy. Hence, the US dollar's purchasing power is defined with respect to US national output, the United Kingdom (UK) pound's purchasing power defines those UK goods and services that can be purchased by the British currency, and so on. In other words, through production each national economy gives its national money a physical content, represented by currently produced goods and services, which are the real object of its purchasing power.

Inflation is therefore the result of an anomaly that affects the initial relationship between money and output established by production. In this sense, inflation is the manifestation of a pathology that increases the number of money units without increasing production (Box 11.1).

Let us expand on this by critiquing the traditional explanations of inflation as presented above. Having defined inflation as a persistent rise in the price level, orthodox economists explain it by referring to the general price level. According to general equilibrium analysis, prices are relative. Goods are exchanged against goods on the commodity market, and it is through their exchange – so the orthodox story goes – that their relative prices are determined. Hence, orthodox economists claim that the price of any given commodity can only be expressed in terms of another commodity against which

BOX 11.1

THE LOGICAL IDENTITY BETWEEN GLOBAL DEMAND AND GLOBAL SUPPLY

To live above one's means cannot increase global demand with respect to global supply. Anyone can increase their expenditures beyond their current income if they obtain a loan on the financial market. Yet, the set of economic agents (considered as a whole) cannot exert a global excess demand, because the excess demand of some agents can only be financed by the equivalent, negative excess demand of other agents. Gifts apart, my expenditures on the products market can exceed my income only if I am a net seller of financial claims: that is, only if somebody else lends me part of their current income. On the whole, excess demands on the products market are necessarily matched by excess supplies on the financial market, so that no inflationary pressure on prices can derive from credit.

the former is exchanged. But how can this conception of prices be reconciled with the undisputable fact that the general price level is expressed in terms of money? Their answer is simply that money itself is essentially a commodity; that is to say, an asset that is exchanged against the goods and services offered in the market. The general price level would then be nothing but the relative price of money in terms of a representative bundle of goods and services defined as a composite commodity. In fact, in the quantity theory of money, the price level is determined by relating the quantity of money to the available quantity of produced goods and services.

Hence, advocates of the quantity theory of money, also known as monetarists, argue that variations in the price level are due to adjustments between two distinct stocks: the quantity of money as determined by monetary authorities (essentially central banks), and the quantity of available output as determined by current production. Inflation would therefore result from an excessive growth of the money supply, which increases the quantity of money without being offset by a corresponding rise of the composite commodity. To avoid an inflationary increase in the general price level, monetary authorities would have to control the growth of the money supply, keeping it in line with variations in the quantity of produced goods and services (Box 11.2).

Modern monetary analysis, which proves the inadequacy of the quantity theory of money and its failure to explain inflation, has shown that this old-fashioned way of looking at money as a stock is wrong. The idea that money can be created as a net asset by central banks is far-fetched to say the least.

BOX 11.2

HOARDING AND DEFLATION

Savings can never take the form of hoarding and be the cause of a deflationary decrease in global demand. Any economic agent is free to postpone at will the final expenditure of their income; however, this has no impact on global demand. Saved-up income is necessarily deposited with banks, and banks lend it at once. The decrease in demand caused by savers is therefore perfectly balanced by an equivalent increase in demand as a result of borrowers' expenditure. This is due to the necessary correspondence between debit and credit as sanctioned by the principle of double-entry bookkeeping, which holds good irrespective of economic agents' behaviour. Analysis shows that income is immediately deposited and lent from the moment of its formation. The nature of money and income is such that hoarding is always and necessarily equal to zero.

Neither human beings nor any of their institutions can create a positive asset out of nothing. Bank money is issued by means of double-entry bookkeeping and can thus only be defined as an asset–liability in accordance with the principle of the necessary balancing of assets and liabilities. If money can nevertheless be endowed with a positive value – its purchasing power – it is because it is associated with production. It is because of its 'integration' with output that money acquires a positive purchasing power. A correct explanation of inflation has to be consistent with the bookkeeping nature of bank money. What has to be investigated is how the strict relationship established by production between money and output can be altered.

Inflation and the price level

A first point that has to be clarified concerns the conventional definition of inflation as a persistent increase in the general price level and its measurement through the Laspeyres formula (see above). In fact, any increase in the price level does not necessarily imply a variation in the relationship between money and output. Indeed, it can easily be shown that: (1) there are persistent increases in prices that do not reduce the purchasing power of money in the least; (2) stability of the general price level can mask the existence of inflation.

The first example is provided by the state's decision to substantially increase indirect taxation, say through an increase in taxes on fuel. The ensuing sharp rise in the price of fuel would certainly have an impact on the purchasing power of many households; yet, would it also reduce the purchasing power of money? The answer is no, because what would be lost by these households would be earned by the state, so that the totality of economic agents (state included) would still enjoy the same purchasing power. The result of the increase in indirect taxation is thus an increase in prices that is not of an inflationary nature. Money would be redistributed among the different categories of economic agents, but its purchasing power would not be reduced (Box 11.3).

Our second example concerns the impact of technological progress on costs and prices. It is a fact that technological improvements generally reduce production costs so that, in the absence of inflation, prices should decrease over time. Let us suppose that, despite the generalized presence of technological or structural improvements, the level of prices remains constant. By identifying inflation with a rise in the general price level, orthodox economists are bound to conclude that there is no inflation, as the price level is the same across the relevant periods. In reality, the correct conclusion would be that it

BOX 11.3

AGENTS' BEHAVIOUR AND INFLATION

Orthodox theories consider economic agents' behaviour to be the root cause of inflation and unemployment. Households' excess demand is thus pinpointed as the cause of inflation, and their excess saving (hoarding) as that of unemployment. What these theories fail to see is that behaviour exerts its effect on an income already formed, that is to say, on the result of production; whereas economic disequilibria are generated by the operation, namely production, that brings about this very result. It is within the mechanism of the payments system characterizing the actual working of any monetary economy of production that one can find the origin of inflation and unemployment.

is due to inflation that the price level, which should have decreased because of the consequences of structural change, has remained constant.

Inflation and the cost of living

Orthodox economists use the consumer price index (CPI) to measure inflation. The CPI is a numerical index (calculated with the Laspeyres formula) attributed to a standardized basket of goods and services that is supposed to include the commodities purchased by a representative household during a given period of time. The use of the CPI has often been criticized by stressing the fact that a comparison between consumer price indexes calculated at different points in time requires the basket to be made up of the same goods in the same proportion. This is obviously not the case, which makes it difficult if not altogether impossible to compare heterogeneous baskets.

But this is not the only criticism that can be levelled against the use of the CPI. The fact that among the goods and services introduced in the representative basket are goods and services produced abroad is another critical element, which shows that the CPI cannot be considered an appropriate measure of the rate of inflation. In fact, foreign goods and services have nothing to do with the definition and determination of the purchasing power of a national currency. Only domestically produced goods and services enter the definition of the output whose production attributes a purchasing power to a country's national currency. It therefore follows that, since variations in the prices of imported goods and services do not bring about any change in the relationship between domestic currency and domestic output, they cannot cause inflation.

While it is true that a significant increase in the prices of imported goods and services may lead to a persistent rise in the CPI, it is also certain that this has nothing to do with inflation. It is therefore clear that the CPI must be considered in fact as a measure of the variation in the cost of living rather than a measure of inflation. A rise in the CPI indeed entails a fall in the standard of living of all those people who cannot easily modify their consumption habits. It could even entail a generalized reduction in the standard of living of every resident. Yet it would be wrong to conclude that it also reduces the purchasing power of the domestic currency as defined by its relationship to domestic output. If the rise in the CPI were caused by a rise in the prices of imported goods and services, this would mean that imports would cost more in terms of exports. Such a change in the terms of trade could be a problem, but it cannot modify the relationship, established by domestic production, between national money and domestic output.

Before investigating how the relationship between money and output can be altered, let us consider the second anomaly examined in this chapter.

Unemployment

Definition

A first, broad and rather obvious definition of unemployment is that of a state of the economy where part of the working population can no longer be employed and has to be laid off. What immediately appears to be a central element in the analysis of unemployment is the determination of its causes. Indeed, it was by distinguishing between its possible causes that Keynes introduced two categories of unemployment – voluntary and involuntary – whose conceptual distinction is of the foremost importance for understanding this major economic disorder.

Voluntary unemployment

Let us immediately clarify that the adjective 'voluntary' shall not be taken in its literal sense. Voluntary unemployment is not that brought about by free choice; it is not merely due to the decision of people not to work. Keynes provides a series of examples that help us to understand that by voluntary unemployment he identifies a side-effect created by the 'normal' working of a capitalist economy.

Frictional unemployment is one of these examples. People who decide to change their working place (either continuing in a given activity or switch-

ing from one to another) or who have to leave it may not immediately find another working place or another job, because of imperfect information or other distortions in the labour market. Unemployment due to industrial structural changes is another example. Because of technological changes, firms are often led to lay off part of their workforce and only very rarely can they immediately find another job.

In what sense is it appropriate to qualify these examples as cases of voluntary unemployment? The choice of the word 'voluntary' is meaningful only if it is taken to define the free choice of an economic system, whose normal workings entail temporary, sectorial adjustments that affect employment. A monetary economy of production based on capital accumulation is a dynamic system, whose changes may have negative, albeit temporary, effects on employment. By choosing such an economic system, in a certain sense we have also chosen to embrace its unintended consequences, and this is why we define the situation in which people are 'frictionally' or 'technologically' out of work as voluntary unemployment.

Now, contrary to what might be thought in the first place, voluntary unemployment is not a serious disorder, and explains only an extremely small fraction of the overall level of unemployment. Indeed, if voluntary unemployment were the only disorder our economic system suffered from, we would have nothing to worry about. This can be better understood by observing that technological evolution often brings with it a possible solution to the unemployment it creates. This is so because technological improvement increases physical productivity, reduces costs, and creates the conditions necessary to reduce the number of working hours per worker without reducing the number of employees. It is true that today this solution is not feasible, but this is because our economic systems remain infected by the two serious pathologies of inflation and involuntary unemployment. Let us turn our attention to this latter form of unemployment.

Involuntary unemployment

Unfortunately, no definition of involuntary unemployment is given in Keynes's work, yet it can be derived from his definition of voluntary unemployment. Since voluntary unemployment is a temporary form of unemployment resulting from the normal imperfections of our economies, involuntary unemployment can only be the persistent form of unemployment caused by the pathological working of our economic systems. The crucial demarcation is that between order and disorder. Any kind of frictional unemployment consistent with an orderly economic system is but an instance of voluntary

unemployment. On the other hand, any form of persistent unemployment due to a structural disorder of the system is a pathology pertaining to involuntary unemployment.

The question that has to be asked now is: what are the criteria by which we can distinguish an orderly from a disorderly economic system? The answer to this question as given by orthodox economics differs substantially from that advocated by modern monetary macroeconomics. Let us first introduce, critically, the orthodox point of view.

Unemployment as the consequence of labour markets' rigidities and changes in consumers' demand

Economists of all schools are unanimous in considering deflation as one of the main causes of unemployment, where deflation is a situation in which demand falls short of supply on the market for produced goods and services, so that part of current output cannot be sold at current prices. Orthodox economists maintain that this situation can be brought about by variations in economic agents' behaviour, for example a reduction in consumption. If, for whatever reason, income holders were to persistently reduce their propensity to consume, the reduction in demand would force firms to produce less, and the ensuing tendency to lay off workers would be offset only if money wages were perfectly elastic downwards. If not, unemployment would set in and the economy could enter a phase of recession.

This traditional, neoclassical analysis of unemployment also emphasizes the role played by workers and their associations. In the presence of deflation or of any other obstacle impeding firms to clear the market (that is, to sell all their output), equilibrium could be preserved, neoclassical economists maintain, if money wages were free to fluctuate. According to their analysis, unemployment is the consequence of workers', trade unions' and states' unwillingness to let nominal wages fall to the level that would preserve full employment.

Besides ethical and socio-political considerations, this analysis has to be rejected on conceptual and analytical grounds. Its main shortcoming is to consider labour as a commodity and to assume that wages are its price, determined on the labour market by the traditional forces of supply and demand. Orthodox economists seem unaware of the fact stated by Smith, Ricardo, Marx and Keynes, among others, that labour is the source of value and that, as such, it cannot itself have a value. Analogously, being at the origin of commodities, it cannot itself be identified as a commodity. Moreover, since money is given a positive value or purchasing power by production, wages

provide the original income in an economic system. A reduction in nominal wages advocated by orthodox economists would only modify the numerical expression of production without touching on the pathology at the origin of involuntary unemployment. Income would be reduced to the same extent as wages, and deflation would essentially remain unchanged.

Orthodox analysis is indeed wrong, since it rests on a wrong conception of bank money, considered as an asset, and on the belief that savings can be hoarded as if money income were a physical quantity that can be taken out of circulation. In reality, the nature of bank money and the principle of double-entry bookkeeping are such that the totality of income necessarily takes the form of bank deposits. Since bank deposits cannot be hoarded, it immediately appears that saving does not reduce the amount of income available in a given banking system. It thus follows that neither labour markets' rigidities nor saving can be the causes of involuntary unemployment.

Towards a monetary macroeconomic analysis of inflation and unemployment

In this section we introduce the elements for a new analysis, emphasizing the existence of a disruptive inconsistency between the nature of money, income and capital on one side, and the way transactions leading to the formation of money, income and capital are entered in banks' ledgers on the other. We will show that, because the present accounting system of national payments cannot consistently distinguish between these three fundamental conceptual entities, a pathology arises that engenders both inflation and deflation.

On money and income

Money is a spontaneous acknowledgement of debt issued by banks. It is because banks issue it by using double-entry bookkeeping that money is an acknowledgement of debt. Yet, its emission would be meaningless if the object of this debt was bound to remain money. In other words, the emission of money acquires all its significance only if it is associated with an economic transaction endowing it with a positive value. The idea that money can be issued as a positive amount of income irrespective of production is thus wrong. It would amount to claiming that banks can create a positive value out of thin air. In reality, money is transformed into income only if it is associated with production, that is, once goods and services are given a monetary form and are thus transformed from physical objects into economic products (see Schmitt, 1960, 1972; Cencini, 2001, 2005). Figure 11.1 represents the result of production.

Figure 11.1
Production as the association between money and output

As shown in Figure 11.1, income is the result of a transaction whereby output becomes the real 'content' or 'object' of money. A positive income is the result of production and constitutes the strictest possible relationship between money and physical output. This means that production generates the income that defines economic output, which is to say that it is through production that money obtains its purchasing power.

From the distinction between money and income we can derive as a first conclusion that economic purchases can be financed only out of positive incomes. Being deprived of any value whatsoever, money as such should never be used to 'finance' a purchase or to cover a financial deficit. The logical distinction between money proper and income should materialize in a distinction between a monetary and a financial department at the banking level, which would guarantee the practical impossibility for money to substitute for income.

Two simple examples show how inflation would arise each time money is issued to 'finance' a purchase or cover a deficit. The first example refers to what would happen if a central bank were to issue money in order to reduce a country's public deficit. The creation of what we could call 'empty' money (that is, a money with no real content derived from production) increases the number of money units available in the country, but leaves the amount of physical output unaltered. As a result, the same output would be distributed over an increased number of money units, leading to the decrease in each money unit's purchasing power (that is, of its real content) (Figure 11.2).

The second example is slightly more complex and refers to the role played by banks as financial intermediaries. The golden rule banks have to follow when lending the income deposited by their customers is that the sum of loans must equal the sum of deposits. Unfortunately, this rule is not enough to avoid the conflation of a monetary emission with banks' financial intermediation.

Figure 11.2 Inflation as the addition of 'empty' money to income generated by production

Table 11.1 Inflation as caused by banks' excessive lending

Bank B			
Assets		Liabilities	
(1) Firm F	100	Income earners (IE)	100
(2) Client C	110	Firm F	110

Let us imagine a commercial bank, B, with a deposit of 100 units of money income, whose object are the goods and services produced by firm F and financially deposited on the assets side of the bank's balance sheet. Let us consider a given period of time and suppose that a client of B asks for a loan of 110 money units in order to purchase, at a price of 110, the goods and services produced by F. Would bank B, consistent with the 'golden rule', be allowed to lend 110 money units to its client C? Apparently no, yet it is enough to write down the bookkeeping entries corresponding first to production and then to the payment that B would carry out on behalf of C and to the benefit of F to realize that the loan of 110 money units would not be incompatible with the 'golden rule' (Table 11.1).

The first entry represents the payment of F's costs of production by the bank and the corresponding formation of 100 units of money income. F is entered on the assets side of B's balance sheet, because it is indebted to the bank following the payment made by B on its behalf. Since F's output is the object of F's debt to B, goods and services are financially owned by the bank. On the liabilities side of B's balance sheet we find the income generated by the payment of F's costs of production and owned by income earners, IE. This means that bank B is indebted to IE and this is so because IE are the owners

of 100 units of money income deposited with B. The deposit is a loan granted by IE to B, whose object is the product that F owes B, so that income earners are the true initial owners of F's output. This cannot come as a surprise, because the value of money income is precisely given by its purchasing power over current output.

The second entry corresponds to the payment of 110 money units carried out by B to the benefit of F on behalf of C. B's client obtains a loan and spends it in order to purchase F's output; this is why the bank's client is entered on the bank ledger's assets side, while F is entered on its liabilities side. Now, it is the entry of 110 money units on the liabilities side of B's balance sheet that justifies the loan of 110 money units granted by B to its client. The 110 money units earned by F define a deposit formed with B, and it is this deposit that covers the loan of 110 money units.

The 'golden rule' is undoubtedly complied with, yet an inflationary gap of 10 money units is formed that brings about a decrease in each money unit's purchasing power. Indeed, the income created by production has remained equal to 100 money units, and this is what the bank should have lent. The 10 money units lent by B on top of the 100 units of money income deposited by income earners have no real 'content' at all: it is 'empty' money that derives its purchasing power from that of the 100 money units associated with production.

The distribution of 100 units of income over 110 money units implies a loss of purchasing power for each unit of money, and the cause of this inflationary decrease is the superposition of a monetary emission on a financial intermediation. Indeed, bank B feeds part of its loan through money creation. Up to 100 money units, its loan is a transfer of income, namely a financial intermediation, while the remaining 10 units result from a creation of money.

This example is far from providing an all-encompassing explanation of inflation. The inflationary gap formed because of B's excess lending is indeed bound to be neutralized when C reimburses its debt (technically, the inflationary gap is not cumulative in time) and banks tend to control their lending activity very strictly in order to avoid interbank indebtedness. Yet, this example shows how a monetary disorder may arise out of a structural lack of respect for the logical and factual distinction between money and income (Box 11.4).

The next fundamental step is to show that inflation and deflation are the twin effects of a process of capital formation in which the essential distinction between money, income and fixed capital goes unrecognized.

BOX 11.4

AGENTS' BEHAVIOUR AND STRUCTURAL–MONETARY REFORMS

The lack of impact of economic agents' behaviour over monetary disorders – to wit, inflation and deflation – is also present in the case of the reform allowing for the passage from capitalism (disorder) to post-capitalism (order). Indeed, the reform of the bookkeeping mechanism of national payments has no impact whatsoever on the behaviour of economic agents, who will be free to adopt whatever decision, socially legitimate, they consider adequate to improve their economic well-being. The causes of economic and financial crises characterizing today's capitalism are structural, in the sense that they are identifiable in the payments system adopted by the economy. The reform calls for the implementation of a new structure enabling the automatic regulation of banks' accounting: it will provide for a new structural framework within which economic agents will carry out their activity in the most liberal way.

The logic of fixed capital formation

As acknowledged by the greatest economists of the past, the first necessary step leading to fixed capital formation is saving. This simply means that in order to have fixed capital goods, part of current income has to be saved and invested in a new production of instrumental goods. If the totality of currently produced income were spent for the final purchase of current output, no fixed capital could be formed and the entire production would consist of consumption goods only. Hence, by subtracting part of current income from consumption, saving creates both a monetary fund that can be invested in the production of instrumental goods, and a stock of consumption goods that will be sold in a later period to income earners.

In any monetary economy of production saving required for the formation of fixed capital takes the form of invested profit. Instead of being spent for the final purchase of current consumption goods, part of current income is transferred to firms and used by them to finance a new production. Now, like any other production, that of instrumental goods leads to the formation of a new income. What happens then to the income transferred as profit to firms and invested by them? The only logical answer coherent with the fact that production is at the origin of income is that it is transformed into fixed capital. The production of consumption goods creates the income required for their final purchase, whereas the production of instrumental goods transforms saved-up income into fixed capital. At the same time, the new income

formed by the production of instrumental goods gives income earners the power to purchase the consumption goods previously stocked by firms.

Let us illustrate this process by a numerical example. Suppose the initial production of consumption goods to be equal to 100 money units. If firms were to sell these goods at the market price of 125 and if income earners were to spend all their income to this effect, firms would obtain 100 money units and give up consumption goods worth 80 money units. The difference between firms' inflows (100 money units) and outflows (goods worth 80 money units), equal to 20 money units, would define their profit. The investment of this profit would have a double effect: it would transform it into a fixed capital equal to 20 money units, and give rise to a new income of 20 money units. Fixed capital would define the monetary face of the newly produced capital goods, while the new income would give IE the right to purchase the 20 money units of consumption goods still unsold by firms.

If banks practised double-entry bookkeeping consistently with this logical process of fixed capital formation, the investment of profit would be entered as shown by Table 11.2.

The first thing to observe is the distinction introduced – by Schmitt (1984) – between a financial department (department II) and a department of fixed capital (department III), the first department (which we do not explicitly introduce here) being the money department. Book-entry (1) results from the formation of a profit, owned by firms, whose real object is the product

Table 11.2 The bookkeeping distinction between income and fixed capital and the investment of profit

Bank's financial department (II)

Assets		Liabilities	
(1) Stock of consumption goods	20	Firms	20
(2) Firms	20	Department III	20
(3) Firms (instrumental goods)	20	Income earners	20
Firms (instrumental goods)	20	Department III	20
Stock of consumption goods	20	Income earners	20

Bank's fixed capital department (III)

Assets		Liabilities	
(2) Department II	20	Firms	20

formed as a stock of still unsold consumption goods. The investment by firms of all their profit, equal to 20 money units, gives rise to book-entries (2) and (3), where entry (2) represents the transformation of profit into fixed capital through its transfer from department II to department III, whereas entry (3) shows the formation of a new income generated by the new production of instrumental goods. As a result of the whole process, income earners own the income necessary to purchase the stock of consumption goods, and firms own the fixed capital deposited in the third department, whose real object is a set of instrumental goods that will be used to increase physical productivity.

The actual process of capital formation

The substantial gap between the logical and the actual process of capital formation lies in the present lack of a distinction between financial and fixed capital departments. Having failed so far to understand the need for a clear and rigorous conceptual and practical distinction between income and capital, economists and bankers keep entering the transactions that we have described in the previous subsection in one and the same banking department. The bookkeeping entries defining the formation of fixed capital that we can find today in the stylized banks' balance sheet are shown in Table 11.3.

As entry (2) shows, the investment of profit corresponds to its expenditure on the factor market. Profit feeds the payment of the costs incurred by firms for the production of instrumental goods. Instead of being transformed into fixed capital and preserved as such in department III, profit is spent on the factor market. The consequences of this expenditure are the following: (1) instrumental goods are lost to income holders and appropriated by what Schmitt (1984) calls 'disembodied firms'; (2) money units obtained by wage earners are deprived of their physical content.

The presence of disembodied firms characterizes the pathological state of our economic system, where the existence of fixed capital denotes a process of alienation whereby the production of instrumental goods leaves income

Table 11.3 The actual bookkeeping entry relative to fixed capital formation

Bank			
Assets		Liabilities	
(1) Stock of consumption goods	20	Firms	20
(2) Firms (instrumental goods)	20	Income earners	20

earners empty-handed. A simple and straightforward question may help to clarify the matter: is it logical for an income transformed into fixed capital to remain available on the financial market? The answer is clearly no. However, this is precisely what happens in the actual system of national payments: the profit invested in the production of instrumental goods reappears as a bank deposit owned by income earners and lent by banks. The result is that fixed capital is spent while it should be saved up in the banks' third department, and money deprived of its 'real' content is lent on the financial market.

In a textbook for undergraduates the analysis can only aim at clarifying the terms of the problem and providing the conceptual elements required for its further development. We will therefore conclude this chapter by summarizing the remaining steps leading to a full explanation of inflation and unemployment:

1. The empty money formed as a result of the investment of profit is clear evidence of the inflationary character of the actual process of fixed capital formation. Yet inflation settles in when instrumental goods are used in the production of new goods and services and have therefore to be amortized.
2. Amortization, in its strict macroeconomic sense, implies the production of amortization goods and modern monetary macroeconomics shows that, when fixed capital is formed pathologically, amortization has two consequences: namely, inflation and overaccumulation of instrumental goods. If fixed capital could accumulate boundlessly, inflation would be a minor disorder, because the loss of money's purchasing power would not impede the substantial rise of our standard of living owing to capital overaccumulation.
3. The problem is that capital has to be remunerated and that this is done out of profits. A constant increase in capital requires a corresponding rise in profits. Now, albeit rising in absolute terms, profits grow at a lower rate than capital, which is why the rate of profit (the ratio between total profit and total capital) has been steadily declining in recent years.
4. The fall in the rate of profit becomes worrisome when it comes to a level close to that of the market rate of interest. From that moment on, national economies have to slow down their capital accumulation, otherwise they would no longer be able to remunerate their capital. The reduction in fixed capital accumulation means that firms must reduce their investment in the production of new instrumental goods. This can be done by investing part of their profit either on the financial market or on the production of new consumption goods. In the first case, a rise in unemployment would be the immediate result.

5. If firms were to invest part of their profit in the production of consumption goods, deflation would settle in. Indeed, the supply of consumption goods on the commodity market would increase without being matched by a rise in the income available to finance their purchase. The production of new consumption goods, in this case being financed by the investment of profit, would give rise to empty money, that is, money with zero purchasing power. The difference between total supply (increased by newly produced consumption goods) and total demand (which remains unaltered) defines a deflationary disorder that will inevitably lead to a further increase in unemployment.
6. The solution consists in a reform of the system of national payments and implies the implementation of a bookkeeping distinction between the departments of money, income and fixed capital.

REFERENCES

Boskin, M.J., E.R. Dulberger and R.J. Gordon et al. (1998), 'Consumer prices, the consumer price index, and the cost of living', *Journal of Economic Perspectives*, **12** (1), 3–26.

Cencini, A. (2001), *Monetary Macroeconomics: A New Approach*, London, UK and New York, USA: Routledge.

Cencini, A. (2005), *Macroeconomic Foundations of Macroeconomics*, London, UK and New York, USA: Routledge.

MacLean, B.K. (2015), 'Quantitative easing', in L.-P. Rochon and S. Rossi (eds), *The Encyclopedia of Central Banking*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 414–16.

Mastromatteo, G. and A. Tedeschi (2015), 'Quantity theory of money', in L.-P. Rochon and S. Rossi (eds), *The Encyclopedia of Central Banking*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 419–22.

Palley, T.I. (1996), *Post Keynesian Economics: Debt, Distribution and the Macro Economy*, London, UK and New York, USA: Macmillan and St Martin's Press.

Rossi, S. (2001/2003), *Money and Inflation: A New Macroeconomic Analysis*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Rowthorn, R.E. (1977), 'Conflict, inflation and money', *Cambridge Journal of Economics*, **1** (3), 215–39.

Schmitt, B. (1960), *La formation du pouvoir d'achat* [*The Formation of Purchasing Power*], Paris: Sirey.

Schmitt, B. (1972), *Macroeconomic Theory: A Fundamental Revision*, Albeuve, Switzerland: Castella.

Schmitt, B. (1975), *Théorie unitaire de la monnaie, nationale et internationale* [*Unitary Theory of Money, National and International*], Albeuve, Switzerland: Castella.

Schmitt, B. (1984), *Inflation, chômage et malformations du capital* [*Inflation, Unemployment and Capital Malformations*], Albeuve, Switzerland and Paris, France: Castella and Economica.

Werner, R.A. (2015), 'Credit easing', in L.-P. Rochon and S. Rossi (eds), *The Encyclopedia of Central Banking*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 121–3.

A PORTRAIT OF BERNARD SCHMITT (1929–2014)

Bernard Schmitt was born in Colmar (France) on 6 November 1929, and died in Beaune (France) on 26 March 2014. After his graduate studies he obtained his PhD in Paris (France) and spent a year at the University of Cambridge (United Kingdom), where he worked with Dennis Robertson, Piero Sraffa and Joan Robinson. He was a Professor of Macroeconomics and Monetary Economics at the Universities of Burgundy (France) and Fribourg (Switzerland) from 1965 to 2000. His research work has been awarded silver and bronze medals by the Centre National de la Recherche Scientifique in France.

It was while writing his doctoral dissertation (Schmitt, 1960) that he had his first intuition about the double-entry nature of bank money and the logical distinction between money and income. Through rigorous analysis he was able to show that, since money's purchasing power is generated by production, the identity between output (macroeconomic supply) and income (macroeconomic demand) is the fundamental law upon which macroeconomics has to be founded. Income and output are the twin results of production, which means that if a difference appears between macroeconomic demand and supply, it can only be of a pathological nature. It is in his 1984 book, *Inflation, chômage et malformations du capital* (*Inflation, Unemployment and Capital Malformations*), that Schmitt provides a full analysis of these two pathologies. In wha t is one of his masterpieces, he shows that today the investment of profit leads to the formation of a fixed capital that is not owned by income holders. This 'expropriation' is caused by a lack of consistency between the actual system of national payments and the logical distinction between money, income and fixed capital. This analysis allowed Schmitt to explain inflation and unemployment as the consequences of fixed capital amortization, and to propose a reform to resolve the inconsistency.

The originality and relevance of Schmitt's analysis lies in a truly macroeconomic conception of macroeconomics and in the development of a quantum monetary theory of production and exchange. The concept of quantum time introduced by Schmitt has no parallel in economics. It establishes macroeconomics as a science, and provides the foundation for a new approach based on absolute instead of relative exchange.

The other field covered by Schmitt's analysis is that of international payments. Since the 1970s he provided a series of writings on the logical shortcomings of the actual (non-)system of international payments and on a critical assessment of European monetary union. Of a particular interest is his 1975 book on the circuits of national and international money, which was followed by numerous post-Keynesian contributions to the so-called theory of the monetary circuit.

Schmitt's long-lasting research on countries' external debts deserves a central place among his most important contributions. Thanks to an in-depth analysis of the role of money in the payment of a country's deficit, he shows that the sovereign debt crisis that so deeply affects our economies at the time of writing is of an entirely pathological origin due to the imperfect understanding of the principles of international transactions. Developed since the 1970s, this analysis is Schmitt's latest legacy and provides both an explanation of and a remedy for the very formation of countries' sovereign debts.

EXAM QUESTIONS

True or false questions

1. Inflation is an increase in the general price level.
2. The central bank originates inflation when it increases the money supply.
3. The conflict theory of inflation explains inflation by an increase in production costs.
4. The cost-push theory of inflation explains that wages and prices on the goods market increase as a result of conflicting claims on income.
5. Inflation can occur even when the price level remains stable.
6. Involuntary unemployment is the result of a structural disorder in the working of the economic system.
7. Inflation and involuntary unemployment stem from the same problem and have nothing to do with agents' behaviour.
8. Banks' double-entry bookkeeping does not yet distinguish money from income.
9. Banks' credit policy cannot originate inflation.
10. Higher liquidity requirements for banks cannot impede inflation.

Multiple choice questions

1. Inflation:
   a) is an increase in the general price level;
   b) is an increase in the cost of living;
   c) is a loss in money's purchasing power;
   d) is the result of a conflict in the distribution of income.
2. Money:
   a) is issued by banks with a positive purchasing power;
   b) is the most liquid financial asset;
   c) can be issued by central banks only;
   d) is an asset–liability issued by any banks.
3. To avoid inflation:
   a) the central bank must control the money supply;
   b) the central bank must guarantee price stability;
   c) a monetary–structural reform of banks' bookkeeping is needed;
   d) banks should become mere financial intermediaries.
4. According to the demand-pull theory of inflation:
   a) inflation results from excessive demand on the goods market;
   b) inflation can be curbed if wages do not increase;
   c) the central bank may support government spending;
   d) public deficits are not a problem.
5. A structural reform of banks' book-keeping:
   a) must respect the essential distinction between money, income and fixed capital;
   b) cannot avoid inflation;
   c) must go along with an increase in banks' liquidity and capital requirements;
   d) could give rise to involuntary unemployment.
6. Which of the following is at the origin of involuntary unemployment?
   a) Workers being dismissed because of a lack of demand on the goods market.
   b) Fixed capital accumulation resulting from firms' investment of profit.
   c) Firms' decision to invest their profits on the financial market.

d) Banks' decision to reduce the provision of credit lines.

7. The Laspeyres price index:
a) underestimates inflation;
b) measures the cost of living;
c) captures technological change;
d) cannot measure inflation properly.

8. The cost-push theory of inflation:
a) explains inflation with regard to banks' behaviour;
b) explains inflation as a result of a conflict in income distribution;
c) considers banks as mere financial intermediaries;
d) none of the above.

9. According to Keynes, voluntary unemployment:
a) stems from the normal working of a market-based economy;
b) stems from workers' decision not to work at the market wage level;
c) cannot be reduced with an increase in public spending;
d) is the result of banks' restrictive credit policies.

10. An orderly-working economic system implies:
a) a bank's bookkeeping structure distinguishing money, income and fixed capital;
b) no involuntary unemployment in the economy as a whole;
c) the absence of disembodied firms;
d) all of the above.

12

The role of fiscal policy

Malcolm Sawyer

OVERVIEW

This chapter:

- explains that fiscal policy relates to the balance between government expenditures and tax revenues;
- argues that fiscal policy should seek to set that balance to influence the overall economy; more specifically, but not exclusively, output and employment;
- points out that the need for fiscal policy arises from the perspective, denied by mainstream economists, that the capitalist economy is subject to cyclical fluctuations and unemployment arising from inadequacy of aggregate demand, as discussed in Chapter 10.

KEYWORDS

- **Automatic stabilizers:** They concern the effects of falling (rising) tax revenues and rising (falling) transfer payments (unemployment benefits, notably) on damping down the effects of a fall (rise) in demand on the level of economic activity.
- **Budget deficit:** It is the difference between government expenditure and tax revenues.
- **Crowding in:** It reflects the post-Keynesian view that an increase in government expenditure leads to a higher level of private demand, stimulating economic activity.
- **Crowding out:** It reflects the mainstream view that an increase in government expenditure, increasing the budget deficit, would reduce investment expenditure (through, for example, upward pressure on interest rates) and other forms of private expenditure.
- **Fiscal policy:** It is the economic policy in respect of the size of the budget deficit and its impact on the level of economic activity.

Why are these topics important?

Fiscal policy is concerned with public government expenditures, on the one side, and tax revenues, on the other side; and more specifically about the balance between them, that is, the overall position of the government budget whether in deficit or surplus. The government budget position is the difference between government expenditures and tax revenues. When expenditures are greater than tax revenues, a government has a deficit, whereas a surplus occurs when government raises more tax revenues than it spends.

Government expenditures can be usefully divided into four categories: (1) current expenditures on goods, services and employment; (2) capital expenditures (often referred to as public investment); (3) transfer payments (social security and welfare payments); and (4) interest on debt. Dividing fiscal expenditures into these four categories is relevant to better understand fiscal policy for the following three reasons. First, both current and capital expenditures are of particular importance, as each of these expenditures leads directly to the use of resources, including labour. Second, these distinctions enable us to set up the primary budget position, which is defined as the difference between government expenditures and tax revenues, but excluding interest payments. Third, the argument is often made that governments should be allowed to borrow for investment purposes but not to cover current expenditures. In doing so, critics try to apply the same analysis to public finances that is discussed with respect to private finances. In other words, they attempt to compare the government to a firm (or a household) that borrows for investment in the hope that the investment will yield future profits. The current budget position is then given by current expenditures plus transfers and interest payments, minus tax revenues. For the purposes of fiscal policy and its impact on the level of economic activity, it is the overall deficit that is relevant.

Is there a need for fiscal policy?

Fiscal policy affects the level of aggregate demand in the economy, and thereby the level of economic activity. In turn there can be a range of reasons (stated or unstated) for concern over the level of aggregate demand: its impacts on the levels of unemployment and employment, the current account position, inflation, and so on. Discussion of fiscal policy has to involve (often implicitly) a theoretical framework of how a capitalist economy operates. A simple illustration of this relates to the question as to whether a market economy is stable or is subject to major fluctuations, and whether there are strong forces within such an economy, which quickly move the economy to full employ-

ment. If the answers to these questions are 'stable' and 'yes', then there is little need for fiscal policy; if the answers are 'lacking stability' and 'no', then there is considerable need for fiscal policy.

I focus here on a closed economy, that is, one that does not engage in international trade, in order to highlight the key issues. There is a basic relationship between the private sector and the public sector, which states:

$$S^a - I^a = G^a - T^a \tag{12.1}$$

where S is private saving, I private investment (the left-hand side of the equation can thus be referred to as net private saving), G is government expenditures and T tax revenues. Government expenditures here include current expenditures, public investment, social transfers and interest payments on public debt. The equation can be viewed as a national income accounts identity when it refers to outcomes. This is done by adding the superscript a after each variable, signifying outcome (often referred to as an *ex post* relationship). This equation indicates that if the private sector lends (private saving exceeds investment) then the public sector borrows, and vice versa.

The relationship between the public sector (the government) and the private sector (firms and households) can also be considered in terms of an equilibrium relationship as in equation (12.2), where the superscript d after a variable means the desired level of that variable. For example, S^d means desired level of savings:

$$S^d - I^d = G^d - T^d \tag{12.2}$$

We may postulate the variables on which desired savings and so on depend. A simple example would be when intentions to save and the intentions to invest both depend on the rate of interest and the level of income, government expenditure is taken as set by the government, and the tax revenue depends on the tax rates set by government and the level of economic activity. Then:

$$S(r,Y) - I(r,Y) = G - t(Y) \tag{12.3}$$

where r is a representative interest rate, Y is income. Hence, in this equation, both savings and investment are viewed as depending on the rate of interest and on income; and $G - t(Y)$ is the government budget position.

The traditional mainstream view

In Figures 12.1a and 12.1b, two representations of equation (12.3) are provided. In the first representation (Figure 12.1a), desired savings and desired investment are each drawn as a function of the rate of interest, with the level of income held constant at the full employment level, with the savings function shown as a positive relationship to the rate of interest, and the investment function as a negative relationship. In other words, at a higher rate of interest, the desire to save would be higher but the desire to invest lower. There is then seen to be an equilibrium interest rate r^* (sometimes referred to as the 'natural rate of interest'), which would ensure the equality between savings and investment at full employment, Y^* (Box 12.1).

In the second representation (Figure 12.1b), savings and investment are now drawn as a function of income. Now, both functions are positively sloped, which means that as income increases, both savings and investment would

Figure 12.1a Portrayal of the 'loanable funds' approach

BOX 12.1

NATURAL RATE OF INTEREST

The 'natural rate of interest' can be understood in terms of the equilibrium rate of interest that arises from equality between the desire to save and the desire to invest. This is illustrated in Figure 12.1a, based on the idea that both savings (positively) and investment (negatively) are influenced by the rate of interest. In that figure, the natural rate of interest would be equivalent to r^*.

Figure 12.1b
Investment and savings linked with income

increase, although at different rates. Given the slopes, it is clear that as income increases, savings would increase faster. Nevertheless, a given r^* would ensure an equilibrium level of income Y^* as indicated.

The key proposition from the mainstream is that full-employment income can be achieved through a suitable choice of the interest rate. In other words, there is only one level of the interest rate, that is the 'natural' level r^*, which guarantees full employment income, Y^*. The general level of interest rates is guided by the policy rate of interest set by the central bank, and the achievement of an interest rate corresponding to r^* would depend on decisions of the central bank. It requires that r^* is indeed positive, as it is generally difficult to generate negative interest rates (the nominal interest rate is generally regarded as subject to a lower bound of zero, and negative real interest rates then require significant inflation). But there is another strong assumption built into this analysis: namely, that the tendencies to save and to invest are broadly similar.

The mainstream view plays down the potential role of fiscal policy. A major reason for this is in effect that fiscal policy (and budget deficits) are not required in order to reach full employment. It can be seen from equation (12.3) that if the left-hand side of the equation is zero, then the right-hand side will be zero. Hence, if intended saving and investment are in balance, then the required public deficit would be zero. The mainstream view has been that intended saving and investment are indeed close to being in balance. This may arise from the interest rate being at the 'natural rate', bringing saving and investment into balance.

The post-Keynesian perspective

For the post-Keynesian view three points stand out. The first is that savings and investment are determined by many factors, of which the rate of interest is only a minor player. Insofar as Figures 12.1a and 12.1b have any validity, the intersection of *S* and *I* may often take place at a negative rate of interest, and in practice a negative rate of interest of any size is difficult to achieve. During the 2010s, the interest rate of central banks came close to zero (less than 1 per cent) and in the euro area the central bank interest rate has fallen at times below zero. In real terms (that is, after allowing for inflation) interest rates became negative, and yet in many countries there were high levels of unemployment.

The second point is that interest rates are not set by what is termed the loanable funds approach (which is in effect represented in Figures 12.1a and 12.1b), but by the actions of the central bank and by liquidity preference (that is, the relative preference for more liquid assets such as money compared to less liquid ones), as discussed in Chapters 5 and 7.

The third point is that saving and investment decisions are undertaken by different people (households and firms in the case of savings, firms in the case of investment) and that the forces influencing these decisions are quite different. There is little reason to think that the general tendencies to save and to invest will be in alignment. Let me illustrate this as follows. The purpose of investment is to make additions to the capital stock in order to produce and sell a higher output (and making higher profits in the process). Investment enables the capital stock to grow broadly in line with the growth of expected demand, and the underlying growth rate of the economy. Investment in the private sector is undertaken in pursuit of profits, and the greater the prospects of profits, the greater the incentive to invest. The profits of a firm provide funds for investment.

Now, consider savings. The savings of households will be strongly influenced by their disposable income, while firms will save out of their profits (referred to as retained earnings). It is generally expected that households will save a relatively small proportion of their wages, whereas firms save a relatively large proportion of their profits.

A crucial question regarding budget deficits and their impact on economic activity is: how would savings and investment intentions respond to government expenditures, tax rates and the size of the budget deficit? In fact, much of the differences between mainstream and post-Keynesian views come precisely from the answers given to this question. According to the main-

BOX 12.2

CROWDING-OUT

An increase in government expenditure may 'crowd out' (that is, lead to a reduction in) private expenditure with the overall result that total expenditure does not rise (and may fall). In turn, the level of economic activity would not rise following an increase in government expenditure. This may occur in a situation close to full employment when the resources that are deployed on public expenditure come from resources that would have been deployed on private expenditure. It may occur if individuals and firms respond to the proposed increase in public expenditure by reducing their consumption plans (through Ricardian equivalence) and their investment plans.

stream, increased budget deficits would lead to 'crowding-out': increased public expenditures would lead to a direct decrease in private expenditures; households would save more (consume less) and firms would invest less in response to an announced intention to increase the budget deficit, so that there would be little if any stimulus to economic activity (Box 12.2).

The post-Keynesian perspective, however, is quite different: an increase in government expenditures can have 'crowding-in' effects. It works like this: increases in public expenditures stimulate investment through their direct effects on the level of demand, and their indirect effects on opportunities for investment. This also stimulates consumer expenditures; as a result, economic activity would increase (Box 12.3).

How can we explain this stark difference between the two approaches? An important part of the mainstream view comes from the idea of what is called the Ricardian equivalence, which can be expressed as follows. A budget deficit

BOX 12.3

CROWDING-IN

An increase in government expenditure (or decrease in tax rates) may 'crowd in' private expenditures. Government expenditure creates jobs and pays higher transfer payments. Incomes then rise and the rise in income leads to higher levels of consumer expenditure and stimulates investment activity. A simple representation of crowding-in comes from the multiplier relationship, which summarizes the increase in income relative to the increase in government expenditure.

involves borrowing and the accumulation of public debt, which in turn has implications for future interest payments on the public debt and the repayment of the principal. In other words, any deficits incurred today must be reimbursed at some point in the future. To pay the deficit, the government may sell bonds, and those who have lent to the government will have acquired assets (bonds) and will benefit from the interest payments on these bonds in the future. But, on the other side, the government now has a liability (bonds) and obligations to pay interest on these bonds for many years. Taxpayers anticipate that there will be future tax obligations to meet these interest payments, and as a result taxpayers will be worse off. They will respond to that by reducing their expenditures today. The full Ricardian equivalence occurs when the taxpayers' response is so strong that they cut back on their consumer expenditure to the extent that it matches the stimulating effect of the increase in public expenditure that had given rise to the higher budget deficit. In the end, the net effect on the economy is nil: there is no multiplier effect.

The mainstream view had been extended even beyond Ricardian equivalence to argue for 'expansionary fiscal consolidation'. This was the idea that a planned reduction in the budget deficit would have such an impact on confidence and interest rates that private expenditure would increase, and do so more than the reduction in public expenditure. Overall expenditures and hence economic activity would expand.

It cannot be denied that a rise in investment expenditures, for example, will lead to a decline in the budget deficit; consider equation (12.1). But the 'expansionary fiscal consolidation' proposition reverses the causality and argues that it is a decline in budget deficits that leads to a rise in investment expenditure, which more than compensates for the decline in public expenditure or the increase in tax rates.

The post-Keynesian proposition, however, is that when there is a revival in investment activity, then the budget deficit declines. The revival in investment activity can come from a variety of sources including a rise in confidence and 'animal spirits', the opening up of new technological opportunities, or the recognition that investment is inherently cyclical. Further, attempts to reduce a budget deficit through austerity programmes may be counterproductive: declines in public investment on infrastructure and declines in employment levels are not conducive to higher investment.

BOX 12.4

AUTOMATIC STABILIZERS

The government expenditure and tax system can operate as an automatic stabilizer in the sense that its presence helps to modify the impact of fluctuations in private demand on economic activity. When private demand and economic activity rise, tax receipts rise and some transfer payments (such as unemployment benefits) fall. The rise in taxes and fall in transfer payments would tend to dampen down demand. Conversely, when demand falls, taxes paid fall and transfer payments rise. The overall effects would be to lessen the rise in economic activity generated by rising private demand, and to modify the fall coming from declining private demand. Government expenditure and tax structure can modify the business cycle but do not eliminate it. The degree of stabilization depends on the nature of the tax system, and will be greater the more progressive is the tax system.

The role of automatic stabilizers

Fiscal policy operates as an automatic stabilizer in that it helps to mitigate the effects of fluctuations in economic activity arising from variations in private demand. When economic activity falls and unemployment rises, then tax revenues (being based on income and sales) decline, and some social transfers such as unemployment insurance benefits rise. The decline in tax revenues and the rise in social transfers help to limit the fall in economic activity.

The scale of budget deficits will move with fluctuations in economic activity. Lower economic activity will widen the budget deficit as tax revenues fall and social transfers rise; higher economic activity will narrow the budget deficit (or increase budget surplus) (Box 12.4).

Estimates suggest that a 1 per cent lower output is associated with a rise in the budget deficit of the order of 0.5 per cent of gross domestic product (GDP). The fall in tax revenues cushions the fall in economic activity (as compared with what would happen without taxation). The degree of cushioning would depend on how tax revenues responded to changes in economic activity; this depends on the progressivity of the tax system. A progressive tax system is one in which, taking the full range of taxes into account, higher-income individuals pay more taxes relative to their income as compared with lower-income individuals. In a progressive tax system, tax revenues would rise faster than income.

While automatic stabilizers can damp down the fluctuations in economic activity, they do not eliminate them. Automatic stabilizers can be augmented by discretionary changes: for example, when there is the prospect of an economic downturn, government expenditures can be increased and tax rates decreased, and the downturn contained or eliminated. The idea that fiscal policy should operate in this way became seen as 'fine-tuning'. As such, it ran into a range of criticisms from mainstream economists.

First, it relies on accurate forecasting of future fluctuations in economic activity, so that changes in fiscal policy can be implemented in response. Second, if fine-tuning was used, it would impact upon economic activity in such a way that the outcome would differ from that forecast, and thereby it becomes impossible to validate the forecasts. Third, there would be also a number of lags: lags in recognizing what is happening to economic activity, lags in decision-making, lags in the implementation of the decision, and lags in the impact of the decisions on economic activity. These lags can mean that the impact of the policy arrives too late. There is, though, some inevitable degree of fine-tuning involved with fiscal policy in the sense that budget decisions are made on an annual basis (and sometimes more frequently) (Box 12.5).

Given the criticism above, two questions arise here. The first is whether it is worthwhile seeking to reduce the degree of economic fluctuations. Some have argued that the social welfare gains from reducing (eliminating) fluctuations are rather small. The simple reason for this thinking is that with fluctuations sometimes output is 'too high' and sometimes 'too low', but they broadly cancel out over the cycle.

In contrast, post-Keynesians argue that not only can fluctuations be substantial (witness the Great Recession starting in 2008) but also that full employment is at most reached at the top of the economic cycle, and attempts at

BOX 12.5

LAGS IN THE APPLICATION OF FISCAL POLICY

There are lags when fiscal policy has effects on the economy after the economic conditions change. There are recognition lags: data on the economy only becomes available with a lag and is often subject to revision. Forecasts on the economy have to be prepared to which policy can respond. There are decision lags: it takes time to consider the data and then make decisions. There are implementation lags: the time between when a policy change is announced and when it impacts on economic decisions.

BOX 12.6

PROGRESSIVE TAX SYSTEM

In a progressive tax system, at the individual level, the amount of taxes a person pays rises more quickly than their income. An income tax system is usually progressive in that low levels of income may not be subject to tax, and the rate of tax becomes higher at higher levels of income. Other taxes are often not progressive: for example, excise duties on tobacco may not be progressive, in that consumption of tobacco products generally does not rise in line with income. In a progressive tax system, when the level of economic activity and incomes rise, tax receipts would rise at a faster rate than income.

limiting the scale of fluctuations would only result in lifting the average levels of output and employment (Box 12.6).

The second question comes from asking whether the automatic stabilizers can be enhanced. As hinted above, a more progressive tax regime would increase the power of the automatic stabilizers: a rise in economic activity would increase tax revenues more substantially, and in the other direction a fall in economic activity would reduce tax revenues substantially and cushion the effects of the fall in demand. A further step could be building in some automatic variations in the tax rates: for instance, social security rates could be linked with the level of economic activity. For example, when the unemployment rate rises by more than a prespecified amount in a quarter, the social security contribution rates would be reduced in the following quarter in order to provide a boost to demand and lower the cost of employing labour.

Functional finance and the post-Keynesian approach to fiscal policy

The post-Keynesian approach to fiscal policy is particularly based on the ideas of Abba Lerner and Michał Kalecki (see the profile of Kalecki at the end of this chapter). It is based on two pillars. The first concerns what should be the purposes of fiscal policy and budget positions. In the words of Lerner (1943, p. 355), we should reject 'the principle of trying to balance the budget over a solar year or any other arbitrary period'. From this point of view, the purpose of fiscal policy is first and foremost to pursue a sustained and high level of economic activity and to move the economy as close as possible to full employment. Budgets should not be balanced for the sake of balancing them. Of particular importance, therefore, is the objective to be set for fiscal policy.

BOX 12.7

FUNCTIONAL FINANCE

The proponents of functional finance reject the idea that the government budget position must be balanced. This is a rejection of the necessity of a balanced budget over a calendar year and over longer time periods such as the business cycle. Instead, the position is advocated that the government budget position should be used to secure policy objectives, usually the achievement of full employment.

Lerner's reference to the time period ('a solar year or any other arbitrary period') is also relevant, because while most economists would accept that in the face of an economic downturn governments should not seek to balance the budget in the year concerned, mainstream economists would generally argue that over the business cycle the budget should be balanced. This can be alternatively expressed, and a common idea now is that the cyclically adjusted budget (often referred to as the structural budget) should be balanced. Irrespective of whether we balance over a year or a business cycle, such policies are harmful to economic activity and employment (Box 12.7).

An example of the view of balancing a budget over a business cycle is the Stability and Growth Pact of the European Monetary Union, which states that the public budget should be balanced over the business cycle with an upper limit of 3 per cent of GDP for the fiscal deficit in any year. In the so-called 'fiscal compact' (Treaty on Stability, Coordination and Governance) agreed to by most member countries of the European Union, this has been amended to read a balanced 'structural budget'.

The second pillar is that an unbalanced budget will generally be required to correspond to full employment. The more usual case would be that a budget deficit will be required in this regard, which from equations (12.1) and (12.2) would correspond to a situation where saving intentions tend to exceed investment. The case where a budget surplus would be required is not ruled out, and of course would correspond to investment intentions exceeding savings. In Figure 12.2 the budget positions for some major countries (and for the countries of the European Union as a whole) since 2000 are given.

Two clear pictures arise from Figure 12.2. First, it is apparent that the budget position varies considerably over time. Much of this variation is related to the business cycle: particularly notice the sharp rise in public deficits around

Source: Calculated from OECD Economic Outlook database

Figure 12.2 Budget deficits as a percentage of GDP in selected countries, 2000–18

2008 and 2009 as the Great Recession struck. Second, the budget position is more often in deficit than it is in surplus: budget deficits are the norm. Indeed, if they were not the norm, then the scale of public debt would tend to zero. The persistence of budget deficits could be interpreted, as is done by the mainstream, as the profligate nature of governments, since a budget deficit means that more public services are being provided than are paid for by tax revenues. The alternative interpretation, however, is that budget deficits are generally required to sustain reasonable levels of economic activity, even though full employment itself was not in general achieved. In the countries featuring in Figure 12.2 there is little (or no) evidence that there was overheating of their economies during the years of public deficits. Indeed, to the contrary, public deficits are larger in years of 'underheating', and fall or become surpluses when the economy booms.

In terms of Figure 12.3 (the foreign sector is omitted for convenience), equilibrium between savings and investment would lead to a level of income Y^.

With saving intentions exceeding investment intentions, a budget deficit would be required to reach full employment income Y^*. The 'functional finance' approach would advocate seeking to set the fiscal policy and budget

Figure 12.3
Illustrating the power of fiscal policy

position of the government such that full-employment income is attained. The budget deficit would be the line *AB*. It can be noted that savings and investment would be higher with a budget deficit than without it. When, say, the desire to invest changes, represented by a shift in the investment schedule in Figure 12.3, then the required budget deficit would also change.

Answering the critics

When fiscal policy is operated from a 'functional finance' perspective, then many of the objections raised against the use of budget deficits fall. The first is the 'crowding-out' argument: that budget deficit and government expenditure crowd out private sector economic activity. However, crowding-out can occur only in situations of full employment: if the economy is operating at full employment, then more employed people in the public sector must necessarily come from the private sector. Society would then have to judge and argue over whether public sector activity is more or less beneficial than private sector activity. Now, full employment is obtained only infrequently, and unemployment is the usual economic condition. In these circumstances, there is no crowding-out: public sector activity in fact 'crowds in' private sector activity. Public sector employment is thus higher; the wages of the public sector employees are spent, and that creates further employment in the private sector.

Another frequently heard argument against a budget deficit is that debt is being accumulated, which will be 'a burden to our children and grandchil-

dren'. The government debt takes the form of bonds paying interest, and these bonds are assets for the members of the public who own them. Thus government debt represents an asset for the private owners of government bonds, and those assets are inherited by our children and grandchildren. The interest payments on government debt are a transfer from taxpayers to bondholders.

A frequently asked question in connection with a budget deficit is: 'Where is the money coming from?' There are two aspects to this question. First, any expenditure, whether by individuals, firms or government, can take place only if there is prior possession of spending power: that is, money. But since the government holds an account with the central bank, it can spend from that account through the issuance of central bank money, which is then spent by the government (see Chapter 5).

The second aspect relates to the funding of the deficit. At the end of an accounting period, the government's accounts (in this regard, similar to any individual's) obey the requirements that expenditures equal receipts plus borrowing. The government borrows in the form of the issue and sale of bonds, and this may change the monetary base (that is, money issued by the central bank, which is the sum of cash and notes held by the public and reserves held by banks with the central bank; see Chapter 5). The consolidated balance sheet of government including the central bank would read as follows:

$$G - T = \Delta B + \Delta M \tag{12.4}$$

where Δ signifies a change in the relevant variable. G is government expenditures, T is total tax receipts, so $G - T$ is the public sector deficit; B is the supply of bonds, and M is the monetary base. Note that the government sells bonds but some of these bonds will be bought by the central bank through open market operations.

Now, can we explain how the public deficit is funded? To answer this question, we must understand what individuals do with their savings. On the private sector side (firms and households), savings are necessarily held in a variety of assets: financial assets, which include deposits with banks and other financial institutions, bonds, or shares in corporations. Much savings will be lent to other parts of the private sector, for instance households are lending to firms, sometimes directly and more often indirectly. But what happens when saving intentions exceed investment intentions? At that point, the saving intentions can only be realized if the government runs a budget deficit, which will generate bonds that households can then purchase.

Indeed, in the absence of a budget deficit there would not be an outlet for the excess savings.

Thus, when a budget deficit is operated along functional finance lines, the conditions in which a budget deficit would be required are precisely the conditions in which there is the potential availability of funds. But these funds can be realized only if there is a budget deficit. This point was well made by Kalecki (1990, p. 360):

> the question of how it is possible to increase government expenditure if expenditure on private investment and personal consumption is cut is answered by the fact that there will always be such an increase in incomes as to create an increase in savings equal to the increase in the budget deficit.

In other words, the public deficit generates the necessary savings to fund it. Thus it cannot be argued that public borrowing will place upward pressure on interest rates.

A budget deficit leads to borrowing and hence to a rise in the outstanding debt of government. This raises two issues. First, there is the possibility of escalating debt; second, the budget deficit could become unsustainable. But there is no reason to believe that either statement is true, and here is why (see the Appendix for the algebra to support this analysis). First, consider a budget deficit relative to GDP of d, which persists; if so, then the ratio of government debt to GDP (labelled b) will tend towards d/g, where g is the nominal rate of economic growth. In other words, if there is a persistent budget deficit of 3 per cent of GDP, this leads to a debt ratio of 60 per cent if the nominal GDP growth rate is 5 per cent; that is, d/g is 0.03/0.05, which is equal to 0.6 (60 per cent).

However, what is relevant for the level of demand and the one that appears in the equations above is the overall budget position, which can be written as $d = pd + ib$, where pd is primary deficit (negative if it is a surplus) and i is the interest rate, so ib is the interest the government pays on its level of debt. Recall that the primary deficit is government expenditures minus tax revenues. Hence, when government expenditures are greater than tax revenues, pd is positive, which means a budget deficit.

Rearranging, we can write that $pd = d - ib$. Recall from above that $d = bg$, so we can write $pd = gb - ib$. This leads us to write $pd = (g - i)b$: there would be a primary deficit only when the growth rate of nominal GDP is greater than the interest rate, and a primary surplus when the rate of interest exceeds the nominal GDP growth rate. When the growth rate and the interest rate are

equal, then the primary budget would be balanced, and in effect the government would be borrowing to cover interest payments.

Second, a sustainable public deficit as just indicated may involve a primary surplus, deficit or balance. The primary budget is the budget excluding interest payments. This could be regarded as a summary statement of what the government spends on goods and services, and the transfer payments relative to the tax revenues raised: the former represents the benefits provided to citizens and the latter the costs involved. This view clearly places no value on the interest payments received but which form an income so far as the bondholders are concerned.

The key point here is that fiscal policy should be set to achieve a high and sustainable level of demand in the economy rather than seeking to achieve a balanced budget. As shown here when fiscal policy is used in that manner, the arguments deployed against fiscal policy and budget deficits are invalid.

Concluding remarks

The role of fiscal policy has been controversial in economic analysis. A mainstream view has played down the need for fiscal policy and budget deficits, largely based on the view that there are strong tendencies for capitalist economies to operate at full employment without any need for government intervention. In contrast, the post-Keynesian perspective is that private demand fluctuates and is in general insufficient to provide full employment. Fiscal policy then has a crucial role in dampening the fluctuations and in providing sufficient demand to achieve high levels of employment. However, while the use of fiscal policy is a necessary condition for high levels of employment, it is not a sufficient condition.

REFERENCES

Kalecki, M. (1944), 'Three ways to full employment', in *The Economics of Full Employment*, Oxford: Blackwell, pp. 39–58.

Kalecki, M. (1972), *Selected Essays on the Economic Growth of the Socialist and the Mixed Economy*, Cambridge, UK: Cambridge University Press.

Kalecki, M. (1976), *Essays on Developing Economies*, Hassocks, UK: Harvester Press.

Kalecki, M. (1990), *Collected Works of Michal Kalecki, Volume 1*, Oxford: Clarendon Press.

Lerner, A. (1943), 'Functional finance and the federal debt', *Social Research*, **10** (1), 38–51.

Appendix

The budget deficit can be written as $D = PD + iB$, where D is the overall deficit, PD is the primary deficit (negative if there is a surplus), i is the interest rate on public debt.

The change in the debt ratio is given by:

$$\frac{d}{dt}\left(\frac{B}{Y}\right)\frac{1}{Y}\frac{dB}{dt} - \frac{B}{Y^2}\frac{dY}{dt} = \frac{1}{Y}D - \frac{B}{Y}\frac{1}{Y}\frac{dY}{dt}$$

where Y is the level of income since the change in debt is equal to the public deficit (including interest payments) and the debt ratio is stable when the change in the ratio is zero, which would imply $d - bg = 0$ and hence $b = d/g$. The overall budget deficit can be written $d = pd + ib$, where pd is the primary deficit (negative if a surplus) and i is interest payment on debt. Then $d = pd + ib$, hence $pd = (g - i)b$: there would be a primary deficit when the nominal GDP growth rate is greater than the interest rate, and a primary surplus when the rate of interest exceeds the nominal GDP growth rate. When the latter rate and the interest rate are equal, then the primary budget would be in balance, and in effect the government would be borrowing to cover interest payments.

A PORTRAIT OF MICHAŁ KALECKI (1899–1970)

Michał Kalecki was born in Lodz, in the Russian-occupied Kingdom of Poland in 1899, and died in Poland in 1970. He was self-taught in economics after training in engineering and mathematics. He obtained his first quasi-academic employment in 1929 at the Research Institute of Business Cycles and Prices in Warsaw. Following the award of a Rockefeller Foundation Fellowship, he travelled to Sweden and England, where he remained for the next ten years, including employment during the Second World War at the Oxford University Institute of Statistics. After working for the International Labour Office in Montreal, Canada, Kalecki was appointed deputy director of a section of the economics department of the United Nations secretariat in New York at the end of 1946, where he stayed until 1954. He returned to Poland in 1955, where he was heavily involved in the debates over the role of decentralization and of workers' councils, the speed of industrialization and the relative size of consumption and investment, and problems of economic development. In the second half of the 1950s, Kalecki was the Chairman of the Commission of Perspective Planning, but his official role was effectively ended in 1960. In the last decade of his life, Kalecki was heavily involved with problems of economic development, including seminars organized at the Academy of Sciences, Warsaw University and the Central School for Planning and Statistics.

Kalecki discovered many of the key ideas of post-Keynesian economics, along with Keynes, and many of their important writings in this respect were during the 1930s. Kalecki developed the ideas that the level of aggregate demand was crucial for the level of economic activity, and that fluctuations in investment were a key ingredient in the generation of cycles in economic activity. He also saw that the market power of firms allowed them to charge prices as a mark-up over wage costs and material costs. Real wages (the ratio of wages to prices) would depend on the degree of market power exercised by firms, and the extent to which there was a higher mark-up of prices over wages. The share of profits would also depend on market power, since profits depend on the difference between revenue (price times quantity) and wages. He introduced the 'principle of increasing risk', whereby the risk of default on a loan rises with the size of the loan (relative to the own capital of the firm), which helps to explain why loans are rationed by banks, as a larger loan would carry a higher risk (for the bank) of default on the loan. His perspective on a capitalist economy was one in which there was often unemployment and spare productive capacity; full employment was in general only achieved at the height of the boom, and for the rest of the business cycle there was considerable unemployment. The capitalist economy was subject to periods of boom and periods of bust, and hence cyclical fluctuations and economic crises. In light of the discovery of the 'principle of effective demand', Kalecki argued that there were economic policy tools, notably fiscal policy, available to overcome unemployment. But the achievement of sustained full employment would encounter major political and social obstacles (Kalecki, 1944).

Kalecki also made substantial contributions to the economics of planning (Kalecki, 1972), and to the analysis of economic and social development (Kalecki, 1976).

EXAM QUESTIONS

True or false questions

1. A budget deficit necessarily causes inflation.
2. A progressive tax system operates to exacerbate economic fluctuations.
3. Public debt is a liability of the state and an asset of the owners of debt.
4. Budget deficit rises when income rises, and falls when income falls.
5. An increase in the rate of income tax would stimulate the economy.
6. Ricardian equivalence is that fiscal policy and monetary policy are equivalent.
7. In a slump, there is not enough tax revenue to finance public expenditure, and taxes must be raised.
8. In a closed economy, the size of the budget deficit is equal to the difference between private savings and private investment.
9. The 'natural rate of interest' is the rate of interest that occurs in nature.
10. The 'natural rate of interest' is defined in terms of the rate of interest that equates the intentions to save with the intentions to invest.

Multiple choice questions

1. 'Crowding-out' means:
   a) an expansion of government expenditure leads to a higher level of employment;
   b) a decrease of government expenditure leads to a decrease of output;
   c) an expansion of government expenditure does not lead to an expansion of output;
   d) an expansion of government expenditure leads to inflation.
2. A primary budget deficit is:
   a) total government expenditure minus tax revenue;
   b) tax revenue minus government expenditure;
   c) tax revenue minus government expenditure on goods and services;
   d) government expenditure other than interest payments on debt minus tax revenue.
3. When the pace of economic activity slows, it is anticipated that:
   a) the government's budget position will move from deficit into surplus;
   b) the government's budget deficit will widen;
   c) the government's budget surplus will grow;
   d) the government's budget position will be unchanged.
4. The operation of automatic stabilizers means:
   a) government activities cause economic fluctuations;
   b) a progressive tax system dampens down but does not eliminate economic fluctuations;
   c) a regressive tax system eliminates economic fluctuations;
   d) a progressive tax system exacerbates economic fluctuations.
5. Ricardian equivalence means:
   a) an increase in government spending encourages private investment;
   b) an increase in government spending leads to decrease in private spending;
   c) an increase in government spending leads to a boom in consumer spending;
   d) an increase in government spending must be avoided.
6. In a closed economy, the difference between private savings and private investment is:
   a) always zero;
   b) equal to the size of the government budget deficit;
   c) equal to the size of the government budget surplus;

d) equal to the increase in the money supply.

7. The key idea of 'functional finance' is:
a) finance has a function;
b) the government budget position should be set to achieve full employment;
c) the government budget must be balanced each year;
d) the government budget must be balanced over the course of the business cycle.

8. In the mainstream view there may be no need for a government budget deficit because:
a) the intentions to save and to invest never coincide;
b) the intentions to save and to invest can be quickly brought into equality;
c) the economy is incapable of reaching full employment whatever the government does;
d) government can only spend what it raises in taxes.

9. In the post-Keynesian view, there is a need for fiscal policy:
a) to enable government to bribe the public through increased government transfers;
b) to use government expenditure and tax balance to offset fluctuations in private expenditure;
c) to increase public ownership of industry;
d) to allow government to reduce taxes.

10. In a progressive tax system:
a) tax revenues fall when national income rises;
b) tax revenues are not related with the level of income;
c) tax revenues rise at a slower rate than income rises;
d) tax revenues rise at a faster rate than income rises.

13

Economic growth and development

Mark Setterfield

OVERVIEW

This chapter:

- reviews the historical record of economic growth; it shows that economic growth is a relatively recent phenomenon that has resulted in divergence between the incomes of fast-growing rich economies and slower-growing poorer economies;
- contrasts supply-led, neoclassical growth with demand-led, Keynesian growth;
- outlines three Keynesian economic growth theories (Harrodian, Kaleckian and Kaldorian); they are shown to differ according to whether investment spending or export demand is the key driver of demand formation and economic growth;
- identifies the properties of Keynesian growth, including the relationship between saving behaviour and economic growth, the effects of income redistribution on growth, the relationship between technical progress and economic growth, and the interaction of supply and demand in the growth process.

KEYWORDS

- **Demand-led growth:** Keynesian macroeconomics suggests that the enlargement of aggregate demand – owing to investment spending or export demand – is a critical determinant of economic growth.
- **Divergence:** The economic growth record shows that rich countries have forged ahead while poor countries have fallen behind, increasing global income inequality in the course of growth.
- **Economic growth:** The process by which real income per capita rises over a protracted interval of time.

- **Structural change:** As economies grow, the composition of output and employment changes, resulting in deindustrialization and the rise of the service sector.
- **Supply-led growth:** Mainstream neoclassical theory credits economic growth to expansion in the availability and productivity of labour and capital.
- **Unsteady growth:** Economic growth fluctuates over time, giving rise to the booms and slumps of the business cycle, for example.
- **Vicious and virtuous circles:** Self-reinforcing patterns of slow or fast growth resulting from the cumulative interaction of demand- and supply-side influences on economic growth.

Why are these topics important?

Economic growth is defined as an increase in real income over a protracted interval of time, say one year. It is usually measured as a percentage rate of change. For example, we might say that the economy grew at 3 per cent in one year, meaning that gross domestic product (GDP) is 3 per cent higher this year than it was at the same time last year.

Economists are especially interested in the expansion of real per capita income; that is, the growth in income per person. When per capita income expands, everyone becomes better off on average in terms of their ability to consume material goods and services (although this may not be the best measure of overall economic well-being; see Chapter 20). Even if everyone is better off on average, however, it need not be the case that every individual is better off (or even no worse off) when per capita income rises. This draws attention to the important connection between economic growth and the distribution of income, between individuals, families and social classes.

'Economic development' is a broader term that is commonly used to refer to the processes of economic growth and structural change in lower-income or less-developed countries, such as Brazil, Mexico or China. Development economists often look at the broader impacts of economic growth on things such as health outcomes, demographics and political institutions. The idea that economic growth involves wholesale social transformation, affecting not just the sectoral composition, technology and institutions of the economy but also aspects of civic life such as the division of time between work and home, or the physical location of a country's population, has a long-standing pedigree in economics. In fact, it can be traced back to the Classical economics of Smith, Malthus, Ricardo and Marx. Since the early twentieth century,

Note: The vertical scale measures per capita GDP in 1990 Geary–Khamis dollars.

Source: Maddison (2008).

Figure 13.1 Economic growth in the very long run

however, the study of economic growth has focused more exclusively on the causes of growth narrowly defined: as sustained increases in the level of real per capita income.

What does economic growth look like in practice? It may be surprising to learn that economic growth, as we know it, is a relatively recent phenomenon, confined to the last few hundred years of human experience. This is illustrated in Figure 13.1, which shows economic growth over the last two millennia in selected countries from five different continents. Notice that prior to the eighteenth century there was little evidence of economic growth anywhere. Since then, however, growth has taken off around the globe; although exactly when and to what extent differs markedly between different countries. For instance, the United Kingdom experienced the first industrial revolution and exhibited signs of growth even before 1800. In countries such as Japan and Brazil, however, economic growth is only really evident during the twentieth century; and whereas Japan has already made up ground on early industrializers such as the United Kingdom, Brazil has not.

These initial observations draw attention to several properties of the historical growth record associated with the significance and unevenness of growth during the post-seventeenth-century growth era. To illustrate this, Table 13.1 shows the growth of income over the past two centuries in representative advanced, middle-income, Latin American and Asian economies.

Table 13.1 Real per capita income (1990 Geary–Khamis dollars), 1820–2007[a]

	1820	1870	1913	1950	1973	1989	2007	Average annual growth rate 1820–2007 (%)	Average annual growth rate 1950–2007 (%)
Advanced economies[b]	1095	1989	3653	5832	12 959	17 835	24 346	1.65	2.10
Middle-income economies[c]	774 (70.63)	1 413 (71.05)	2 357 (64.51)	3 592 (61.59)	7641 (58.96)	9768 (54.77)	13 696 (56.25)	1.53	1.97
Latin American economies[d]	700 (63.88)	997 (50.12)	2 332 (63.84)	3 173 (54.41)	5 432 (41.92)	5 981 (33.54)	9 367 (38.47)	1.38	1.59
Asian economies[e]	579 (52.87)	559 (28.10)	740 (20.26)	697 (11.94)	1 659 (12.80)	3 938 (22.08)	9 108 (37.41)	1.47	3.78

Notes:

a. Figures in parentheses are percentages of advanced economies' real per capita GDP.

b. Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Sweden, Switzerland, United Kingdom, United States. Based on Maddison (1991, Table 1.1, pp. 6–7).

c. Former Czechoslovakia, Hungary, New Zealand, Portugal, Spain, Former USSR. Based on Maddison (1991, Table 1.5, pp. 24–5).

d. Argentina, Brazil, Chile, Mexico. Based on Maddison (1991, Table 1.5, pp. 24–5).

e. Bangladesh, China, India, Indonesia, South Korea, Taiwan. Based on Maddison (1991, Table 1.5, pp. 24–5).

Source: Author's calculations based on Maddison (2008).

To give some idea of the significance of the growth rates reported in the last two columns of Table 13.1, consider the average annual rate of economic growth since 1820 in the advanced economies (those that currently have the highest per capita incomes and that, in general, also have the longest histories of economic growth). This growth rate is sufficient to ensure the doubling of real per capita income roughly every 40 years; that is, within the working lifetime of each generation. Since 1820, economic growth has made the average person living in an advanced economy roughly twice as well-off as their parents.

Table 13.1 also shows the unevenness of the growth experience, which has meant that growth in different economies has been marked by processes of forging ahead, catching up and falling behind. Note, for example, that the advanced economies were already the richest economies in 1820, following which they then grew faster than any other group of economies over the next two centuries. As a result, they were even richer (in both absolute and relative terms) by comparison with the rest of the world by 2007. This is reflected in the figures in Table 13.1 that report the real per capita income of middle-income economies, of Latin American economies and of Asian economies as a percentage of advanced-economy real per capita income. Without exception these figures fall over the period 1820–2007 as a whole. The unevenness of growth is also illustrated in Figure 13.2, which shows the gap between the advanced economies and the rest of the world widening from 1820 to 2007. Figure 13.2 uses a logarithmic scale so that a

Note: The vertical scale measures natural logarithm of real per-capita income.

Source: Author's calculations based on Maddison (2008).

Figure 13.2 Forging ahead, catching up and falling behind: the general experience

Note: The vertical scale measures natural logarithm of real per-capita income.

Source: Author's calculations based on Maddison (2008).

Figure 13.3 Forging ahead, catching up, and falling behind: country-specific experiences

change in the vertical distance between any two lines in the figure represents a change in the relative income of the countries represented by these lines.

Over the past two centuries, then, the general growth experience has been one of divergence. The early start and relatively rapid growth of the advanced economies has seen them forge ahead, while the experience of the rest of the world has involved a process of falling behind.

But Table 13.1 and Figure 13.2 also show some evidence of catching up. The final column of Table 13.1 shows that since 1950 the Asian economies have grown almost twice as quickly as the advanced economies. As can be seen in Figure 13.2, this has closed the relative per capita income gap between these economies since the middle of the twentieth century. Figure 13.3 shows that this catching-up process is particularly evident in a small number of very successful economies. A good example is South Korea, which has grown so rapidly since the 1950s that it has now essentially joined the club of advanced economies that enjoy the highest per capita incomes.

Two other features of the historical growth record that are worth emphasizing are the tendency of growth to be unbalanced and unsteady. Growth is unbalanced in the sense that different sectors of the economy grow at different rates, with the result that economies experience what is called

structural change. The most obvious symptom of this process is the changing composition of employment. Over the last two centuries, all of the advanced economies in Table 13.1 have experienced a continuous decline in the share of the workforce employed in agriculture, and a continuous increase in the employment share of the service sector. These trends have been accompanied by first a rise and then, since the middle of the twentieth century, a decline in the employment share of the manufacturing sector. This is called deindustrialization. There is considerable debate about the causes of deindustrialization. One theory suggests that it is the inevitable result of a 'normal' transformation that accompanies economic growth once a country reaches a particular (relatively high) level of real per capita income (the maturity hypothesis). Another theory suggests that, at least in some instances, deindustrialization results from lack of competitiveness and/or bad economic policy, such as persistently overvalued exchange rates (the failure hypothesis).

Economic growth is said to be unsteady because, rather than being constant, GDP growth rates fluctuate over time. Some economists argue that these fluctuations are quite regular and take the form of cycles. The existence of a Juglar or business cycle is widely accepted, but cycles of longer duration such as Kuznets swings (over periods of 25–30 years) or Kondratieff waves (over periods of 40–50 years) have also been proposed. This means that over, say, 50 years the economy will experience a long (20–25 years) boom period of fast growth, followed by a similarly long slump period marked by slower growth, before again experiencing a long boom of faster growth, and so on. Other economists argue that fluctuations of longer duration exist, but that these represent irregular and historically specific episodes or phases of growth rather than a well-defined cycle (see, for example, Maddison, 1991; Cornwall and Cornwall, 2001).

The traditional mainstream view: supply-led growth

What explains the patterns of growth uncovered by this review of the historical record? Answering this question is the purpose of growth theory. As in so many matters of economic theory, economists are divided as to the main causes of economic growth. The most important division is between the mainstream neoclassical supply-side view, and the alternative Keynesian demand-side view.

According to mainstream neoclassical analysis, economic growth is a supply-side process. In other words, it results from the increased availability and/or productivity of the factors of production (such as labour or capital goods),

BOX 13.1

POTENTIAL OUTPUT

'Potential output' is the term used to describe the maximum output that can be produced in the economy at any point in time, given the availability and productivity of the factors of production. Potential output has important implications for the growth of output, by imposing a ceiling on the output path of a growing economy that is related to resource constraints. In mainstream neoclassical growth theory, the economy's potential output path describes the actual output path of the economy, doubling up as the equilibrium rate of growth. In Keynesian growth theory, however, it describes only a constraint on the demand-led actual output path, which must always lie on or below the path traced out by the growth of potential output.

which determine expansion of the economy's potential output. According to Stern (1991, p. 123), 'the study of growth . . . is about the accumulation of physical capital, the progress of skills, ideas and innovation, the growth of population, how factors are used, combined and managed and so on. It is therefore, principally, about the supply side.' In this view, there is no room for the demand side to play a significant role in determining the growth rate. Say's Law rules supreme: 'supply creates its own demand', so demand simply adjusts to match the available supply of goods as technical progress, and the increased availability of capital and labour cause potential output to expand over time (Box 13.1).

Figure 13.4(a) illustrates the mainstream view using the concept of a production possibility frontier (PPF). The PPF shows what combination of any two goods the economy can produce when it fully utilizes all of its available resources.

In the mainstream neoclassical growth process, increases in the quantity and/or productivity of factors of production (labour or capital goods) push the PPF out, from PPF_1 to PPF_2 and PPF_3. Actual output then expands in tandem with these developments (from A to B to C), thanks to the operation of Say's Law ('supply creates its own demand'). In other words, increases in the availability and/or productivity of labour or capital goods means that we both can and do produce more goods of all kinds.

Figure 13.5 illustrates a typical growth trajectory produced by the neoclassical growth process just described.

Figure 13.4
Neoclassical and Keynesian growth processes

Let us assume that the economy progresses along its potential output path (denoted as $\ln y_p$) until time period t, at which point output is reduced by a recession. According to neoclassical growth theory, output recovers quickly and automatically to its potential level (path A), so that by period $t + n$ the economy is once again on its potential output path. Unless its progress is interrupted by another recession, the economy will continue to expand along its potential output path, as determined by the increasing availability and productivity of capital and labour.

The Keynesian view: demand-led growth

Keynesian economists reject Say's Law, arguing that there is no mechanism in the economy that causes aggregate demand to automatically adjust to the level required by potential output. The substance of this proposition

Figure 13.5
Neoclassical and Keynesian growth trajectories

in the short run has been discussed earlier in this book (see Chapter 10). For Keynesians, the same issues and concerns carry over to the longer term. The result is that economic growth is understood to depend on the growth and development of aggregate demand over time. This is called demand-led growth.

The Keynesian growth process is illustrated in Figure 13.4(b). As in the neoclassical case described earlier, increases in the quantity and/or productivity of capital and labour push the PPF out from PPF_1 to PPF_2 and PPF_3. This time, however, actual (as opposed to potential) output expands independently of these developments, from A to B to C. This is because aggregate demand – not the availability and productivity of capital and labour – explains the level of output at any point in time and, as a result, the expansion of output (economic growth) over time. In the real world, productive resources will be chronically underutilized as a result of the Keynesian growth process. The economy will almost always operate in the interior of its PPF (that is, at an actual level of output below potential output), so there will always be unemployed labour and idle capital.

Figure 13.5 illustrates typical growth trajectories produced by the Keynesian growth process. Let us assume once again that the economy initially progresses along its potential output path (denoted as ln y_p) until time period *t*, at which point output is reduced by a recession. In the Keynesian case, many possible trend expansion paths are possible after period *t*. For example, the economy may begin a slow adjustment back towards potential output (path B). Other things being equal, this would mean that the potential output path is eventually regained. However,

the economy may expand at or even below potential (along path C, for example). In this case it will experience what is known as secular stagnation: a prolonged period of slow growth coupled with high or even rising unemployment. Given these two possibilities, one important conclusion is that, in the Keynesian case, the growth trajectory after period t in Figure 13.4 is not predetermined by the economy's potential output. There is no guaranteed return to the potential output path by $t + n$ (or any other point in time thereafter).

Sources of demand in the long run

If the Keynesian growth process is demand-driven, what are the sources of aggregate demand that determine the pace of economic growth in the longer term? Historically, these forces have been idiosyncratic. The United States (US) economy provides a good example of this. Following the Great Depression, almost a decade of weak economic growth ensued during the 1930s, before massive rearmament in response to the advent of World War II eventually boosted demand to the point of fully employing all available resources by the early 1940s. Immediately after the war, the United States experienced a long economic expansion that was propelled and sustained by large-scale infrastructure projects (in particular, the construction of the interstate highways), military spending as a result of the Cold War, and the expansion of international trade. Demand growth faltered during the 1970s and 1980s, partly as a result of deliberate fiscal and monetary policies designed to reduce inflation. It recovered strongly in the mid-1990s, however, based first on a new wave of investment spending on information and communication technologies following the commercialization of the Internet, and then by residential construction and a debt-financed consumer spending boom. In 2008 the dynamics of this demand-generating process unravelled as a result of the financial crisis and Great Recession. Since then, demand formation and (as a result) the US economy have embarked on a long but slow recovery that has left many economists worrying about secular stagnation. (See Cynamon et al., 2013, pp. 16–19 and Harvey, 2014 for more extensive accounts of the US experience of demand formation since the early twentieth century.)

Demand formation in Europe has followed similar patterns to those observed in the United States since the early twentieth century, although once again the precise forces responsible for the expansion of demand have been idiosyncratic. Demand growth after 1945, for example, was propelled in large measure by reconstruction following World War II, with an important supporting role played by the large-scale expansion of the welfare state. More

recently, meanwhile, austerity policies throughout Europe have focused on cutting public expenditures in response to rising public sector deficits and debt. These policies impeded demand formation and economic growth in the aftermath of the Great Recession.

Beyond these historical observations, Keynesian macroeconomic theory suggests that the main drivers of aggregate demand are internal and external sources of autonomous demand. In other words, either investment or exports can usually be seen as the basic cause of demand formation and economic growth. The rest of this chapter will refer to these as the Harrodian/Kaleckian and Kaldorian views, respectively, named after three famous Keynesian economists (Roy Harrod, Michał Kalecki, and Nicholas Kaldor) with whom they can be associated. These views, and their implications for economic growth, can be understood by building on the income determination analysis used to describe the short run (see Chapter 10).

Before proceeding, it is important to note that the emphasis throughout the rest of this chapter is on aggregate demand formation and the key sources of autonomous demand responsible for this. But sustainable growth and development may also require a certain composition of demand. For example, sufficient investment in specific types of capital (private, public, environmental and social), rather than just sufficient investment in total, may be needed. These and other themes related to sustainable growth are discussed in detail in Chapter 20.

Keynesian growth theory

Harrodian growth theory

In Harrodian growth theory, investment by firms is the key source of autonomous demand that drives growth. The Harrodian view can be developed as follows. Let us begin with three equations describing investment by firms, saving by households, and the relationship between saving and investment necessary for macroeconomic equilibrium (equality of aggregate supply and demand) in a closed economy with no fiscally active government sector:

$$I = \beta Y_{-1} \tag{13.1}$$

$$S = s_{\pi} \pi Y \tag{13.2}$$

$$S = I \tag{13.3}$$

In equations (13.1), (13.2) and (13.3), I denotes investment, Y is total output, S is total saving by households, and π is the profit share of income. Equation (13.1) is an investment function, which says that investment increases (with a lag) in response to rising output. This is known as the accelerator effect. It is based on the idea that firms invest so as to make sure that their productive capacity keeps pace with economic expansion. Equation (13.2) describes household saving behaviour. Note that saving takes place out of profit income only (at a rate determined by the propensity to save out of profits, s_π): workers are assumed to spend all of their (wage) income, so they do not save. This simplifying assumption captures the idea that wage earners are generally among the less affluent households in the economy (associated with the bottom 80 per cent of the size distribution of income), and therefore have the highest marginal propensities to spend. Equation (13.3) states the condition necessary for macroeconomic equilibrium.

Combining equations (13.1), (13.2) and (13.3), we get:

$$s_\pi \pi Y = \beta Y_{-1}$$

and solving this expression for Y gives us:

$$Y = \frac{\beta}{s_\pi \pi} Y_{-1} \tag{13.4}$$

By subtracting Y_{-1} from both sides of equation (13.4), we arrive at the following expression for the change in output over time; which, in this simple framework, we can associate with economic growth:

$$\Delta Y = \left(\frac{\beta}{s_\pi \pi} - 1\right) Y_{-1} \tag{13.5}$$

Now suppose that $Y_{-1} = Y_0$ and that $\Delta Y > 0$. This means that we will observe $Y_1 > Y_0$, which according to equation (13.5) will cause ΔY to increase subsequently. Alternatively, if $Y_{-1} = Y_0$ and $\Delta Y < 0$, then we will find that $Y_1 < Y_0$ and ΔY will now decrease from one period to the next. This suggests that investment-driven economic growth is unsteady. As noted earlier in this chapter, the rate of economic growth is unsteady if it varies over time, increasing or decreasing from period to period rather than remaining constant. Note that if the economy alternates between increasing and decreasing growth, then economic growth will fluctuate in a cyclical fashion. The economy will then experience booms and slumps marked by faster and slower growth,

respectively, giving rise to the business cycles, Kuznets swings and/or Kondratieff waves described earlier.

Kaleckian growth theory

Kaleckian growth theory is also associated with the idea that investment by firms is the main driver of demand formation and economic growth. The Kaleckian view can be understood by once again beginning with three equations that describe investment by firms, saving by households, and the relationship between saving and investment necessary for equilibrium:

$$S = s_{\pi}\pi Y \tag{13.2}$$

$$S = I \tag{13.3}$$

$$I = \alpha + \beta Y + \gamma\pi Y \tag{13.6}$$

The difference between the Harrodian and Kaleckian models boils down to the investment equation (13.6), which now suggests that investment increases in response to rising output, Y, and rising profits, πY. According to equation (13.6), investment can also vary independently of either output or profits owing to variations in α. This captures the influence on investment spending of 'animal spirits', that is, the willingness of firms to act spontaneously in the face of fundamental uncertainty about the future (see Chapter 10). Equations (13.2) and (13.3) once again describe household saving behaviour and the conditions necessary for macroeconomic equilibrium, respectively.

Combining equations (13.2), (13.3) and (13.6), we now arrive at:

$$s_{\pi}\pi Y = \alpha + \beta Y + \gamma\pi Y$$

and solving this expression for Y results in the following expression:

$$Y = \frac{\alpha}{(s_{\pi} - \gamma)\pi - \beta} \tag{13.7}$$

It follows from equation (13.7) that the change in output over time – and hence economic growth – can be described as follows:

$$\Delta Y = \frac{\Delta\alpha}{(s_{\pi} - \gamma)\pi - \beta} \tag{13.8}$$

Equation (13.8) tells us that changes in 'animal spirits' – the exogenous determinant of investment spending in equation (13.6) – drive the expansion of aggregate demand and hence output. Unlike the Harrodian case discussed previously, ΔY will not automatically increase or decrease over time. If, however, the size of $\Delta\alpha$ varies, then so will ΔY, and the pace of economic growth will once again fluctuate.

Kaldorian growth theory

One of the hallmarks of Kaldorian growth theory is that it is based on open-economy analysis. As a result, it identifies export demand as the critical source of autonomous demand driving economic growth. The Kaldorian view can be developed by first extending equation (13.3) to allow for international trade. In this case, the macroeconomic equilibrium condition can be stated as follows:

$$S + M = I + Y \tag{13.9}$$

where M and X denote imports and exports, respectively. Imports are added to savings because they represent an additional 'leakage' from the circular flow of income (that is, a reduction in spending on domestically produced goods and services). Exports are added to investment, meanwhile, because they represent an additional injection into the circular flow of income (that is, an increase in spending on domestically produced goods and services).

Now suppose that imports depend on the level of income:

$$M = mY \tag{13.10}$$

where m is the propensity to import. This suggests that as our income rises, we will purchase more goods produced abroad.

If we keep the descriptions of saving and investment behaviour in equations (13.1) and (13.2), then substituting equations (13.1), (13.2) and (13.10) into equation (13.9) gives us the following equation:

$$s_{\pi}\pi Y + mY = \beta Y + X$$

Notice that the precise form of equation (13.1) used here is actually $I = \beta Y$, which mirrors the relationship between I and Y in Kaleckian growth theory (see the second term on the right-hand side of equation 13.6). Solving the equation above for Y produces the following expression:

$$Y = \frac{X}{s_\pi \pi + m - \beta} \tag{13.11}$$

It follows from equation (13.11) that the change in output over time – which we can once again associate with economic growth – can be described as follows:

$$\Delta Y = \frac{\Delta X}{s_\pi \pi + m - \beta} \tag{13.12}$$

Equation (13.12) tells us that the expansion of export demand drives the expansion of aggregate demand and hence output. If we introduce Kaldor's assumption that the saving-to-income and investment-to-income ratios ($s_\pi \pi$ and β, respectively) are roughly equal (see Palumbo, 2009), then equation (13.12) becomes:

$$\Delta Y = \frac{\Delta X}{m} \tag{13.13}$$

The focus on the right-hand side of equation (13.13) is now exclusively on features of international trade (m and X) as determinants of the expansion of output. This is representative of Kaldorian balance-of-payments constrained growth theory, which is explored in detail in Chapter 16.

Summary

Table 13.2 summarizes the main features of each of the three strands of Keynesian growth theory – Harrodian, Kaleckian and Kaldorian – described in detail above.

Table 13.2 A summary of the main features of Keynesian growth theory

	Key equation	Main driver of growth	Nature of growth outcomes
Harrodian model	Equation (13.5) $\Delta Y = \left(\frac{\beta}{s_\pi \pi} - 1\right) Y_{-1}$	Investment	Unsteady (growth will vary from period to period, possibly producing cycles)
Kaleckian model	Equation (13.8) $\Delta Y = \frac{\Delta \alpha}{(s_\pi - \gamma)\pi - \beta}$	Investment	Steady (constant rate of growth)
Kaldorian model	Equation (13.13) $\Delta Y = \frac{\Delta X}{m}$	Exports	Steady (constant rate of growth)

Properties of Keynesian growth theory

The three models presented in the previous section all share important features. One of the key features of short-run Keynesian analysis is the paradox of thrift: that is, the proposition that an increase in the saving rate will depress aggregate demand and hence income and, in the process, leave total saving unchanged (see Chapter 10). The analysis in the previous section reveals that, according to Keynesian growth theory, this result has a counterpart in the longer term, when an increase in the saving rate will depress demand formation and hence growth. To see this, notice that any increase in the saving rate, s_π, causes the terms on the right-hand sides of equations (13.5), (13.8) and (13.12) to become smaller. This, in turn, reduces the rates of expansion of income on the left-hand sides of these equations. The rationale for this result is that the higher saving rate reduces the increments in consumption spending that are associated with any autonomous demand stimulus (regardless of its source), which reduces the rate of expansion of income. What this suggests is that thrift is not a virtue if the objective of society is to raise the rate of economic growth. This result has contemporary relevance, because some Keynesian economists believe that prior to the onset of the Great Recession, falling saving rates in the United States (which according to the paradox of thrift will increase the rate of economic growth) helped to offset the negative growth effects of rising income inequality (see Palley, 2002). In other words, economic growth in the United States during the 1990s and 2000s was higher than it would otherwise have been, because of the effects of the paradox of thrift.

A second key feature of Keynesian growth theory is the relationship between distribution and economic growth. The examination of the historical growth record has already revealed that economic growth is accompanied by changes in the distribution of income. The pattern of income divergence that results from forging ahead and falling behind means that, on an international scale, the rich get richer both absolutely and relatively.

Keynesian growth theory goes one step further, however, suggesting that changes in the distribution of income can be a cause of variations in economic growth. A particular concern is that redistribution of income away from wages and towards profits, which broadly speaking involves redistribution from less-affluent to better-off households (see Atkinson, 2009; Glyn, 2009), will retard the rate of economic growth. This concern is clearly reflected in the Keynesian growth theories presented in the previous section. Notice that any increase in the profit share, π, causes the terms on the right-hand sides of equations (13.5), (13.8) and (13.12) to become smaller. This,

once again, will reduce the rates of expansion of income on the left-hand sides of these equations.

The result just described is similar to that obtained in the short run, when falling wages depress aggregate demand, and hence output and employment. The rationale for the result is that the rising profit share boosts profits and so deprives high-spending working households of income, which depresses total consumption. This adverse effect on demand formation more than outweighs any accompanying positive effect, such as the boost to investment spending caused by a rise in the profit share that is evident in equation (13.6) of the Kaleckian growth model. The net result is that the rate of economic growth declines. The effect of income redistribution on economic growth is a particularly prominent theme in Kaleckian growth theory. Its contemporary importance is reflected in widespread concerns that in economies such as the United States, the negative effects on economic growth of increasing income inequality were – at least prior to the Great Recession – offset by unsustainable increases in household borrowing. This draws attention to the links between economic growth, finance (see Chapter 18) and crisis (see Chapter 19). It also suggests that reduced inequality – which may be an important economic objective in itself – is also conducive to enhancing both the extent and the sustainability of economic growth.

A final important theme in Keynesian growth theory is technical progress. That economic growth is accompanied by technical progress is not in doubt: just think of the many improvements in technology, from the steam engine to the smartphone, that have accompanied the rapid expansion of real per capita income during the growth era described earlier in this chapter. But Keynesian growth theory has always maintained that technical change does not simply accompany, but is instead induced by, economic growth. In other words, technical change is endogenous to economic growth. In Keynesian growth theory, this idea has long found expression in the form of 'Verdoorn's Law', which can be stated as follows:

$$\Delta q = \delta \Delta Y_{-1} \tag{13.14}$$

According to equation (13.14), the expansion of output causes a subsequent expansion of productivity (output per person employed), q. Various reasons have been advanced for this relationship. The first can be traced back to Adam Smith's dictum that 'the division of labour depends on the extent of the market'. In other words, the expansion of output is conducive to specialization by trade: different producers can become expert in the production of a single good or service once the market for that good or service becomes

BOX 13.2

VERDOORN'S LAW

Verdoorn's Law takes its name from the Dutch economist Petrus Johannes Verdoorn. It describes Verdoorn's (1949) empirical finding of a positive relationship between the growth rates of labour productivity and output in manufacturing across countries. Verdoorn's Law is usually written as follows:

$$\frac{\Delta q}{q} = a + b\frac{\Delta Y}{Y}$$

where $\Delta q/q$ and $\Delta Y/Y$ are the proportional rates of growth of productivity (output per person) and output, respectively. The coefficient b is known as the 'Verdoorn coefficient' and the coefficient a measures any exogenous influences on the productivity growth rate. In most empirical studies, the Verdoorn coefficient is found to have a value of about 0.5. This means that a one percentage point increase in the growth rate of output leads to a one-half percentage point increase in productivity growth. Verdoorn's (1949) empirical study, published in Italian, went unnoticed by most economists until it was championed by Nicholas Kaldor. For this reason, Verdoorn's Law is sometimes referred to as the Kaldor–Verdoorn effect.

sufficiently large. A second explanation for Verdoorn's Law is that increased output permits profitable utilization of capital equipment that is more productive but also both expensive and 'lumpy' (that is, difficult to buy in small quantities). Think, for example, of how many vehicles a factory would need to produce in order to justify investment in integrated assembly-line technology of the sort used by the world's largest vehicle manufacturers. A final explanation concerns 'learning by doing': the simple principle that 'practice makes perfect'. The idea is that increases in output increase the amount of 'doing' and, as a result, increase learning by doing and hence productivity (Box 13.2).

One important feature of equation (13.14) is that it shows how the demand-led expansion of actual output affects productivity and hence the economy's potential output (which, as discussed earlier, depends in part on the productivity of the factors of production). Potential output is usually thought of as a supply-side variable. But equation (13.14) suggests that it is influenced by the process of aggregate demand formation that drives economic growth. This demonstrates that, from a Keynesian perspective, the demand and supply sides of the economy (represented here by actual and potential output) are interdependent in the process of economic growth.

According to Kaldorian growth theory, this last insight can be carried one step further. Suppose, then, that in addition to equation (13.14) we also observe:

$$\Delta X = \eta \Delta q \tag{13.15}$$

According to equation (13.15), increased productivity increases export demand, by lowering the costs of production and hence the prices of exports, or by improving the quality of exported goods, or both. Notice that increased productivity now stimulates demand formation (in equation 13.15) causing actual output to expand (in equation 13.13). In other words, the expansion of potential output (as a result of Δq) now influences the expansion of actual output. Putting the various pieces of the analysis together, this means that the expansion of demand stimulates the expansion of actual output (in equation 13.13), which stimulates factor productivity (in equation 13.14) and hence potential output, which further stimulates the expansion of demand (in equation 13.15) and hence actual output (in equation 13.13), and so on. We have now discovered the possibility of self-reinforcing virtuous and vicious circles of economic growth, which can be summarized by substituting equation (13.14) into equation (13.15) and then substituting the result into equation (13.13). This yields:

$$\Delta Y = \frac{\delta \eta}{m} \Delta Y_{-1} \tag{13.16}$$

According to equation (13.16) an initially low rate of economic growth will propagate a subsequently low rate of growth (a vicious circle), whereas if an economy begins with a high rate of growth, its rate of growth will tend to remain high thereafter (a virtuous circle). These outcomes suggest that the growth process is historical in nature, or path dependent: the rate of growth today depends on the rate at which an economy has grown in the past (Box 13.3).

Notice the important relationship between path dependence and the observations made at the start of the chapter about income divergence due to forging ahead and falling behind. As previously noted, over the last two centuries the initially prosperous and rapidly expanding advanced economies have tended to maintain their rapid expansion, while less-prosperous and slower-growing economies in the rest of the world have, in general, remained growth laggards. As a result, the real per capita income of the advanced economies has increased both absolutely and relative to the rest of the world. According to Kaldorian growth theory, these patterns of forging ahead and falling behind in the economic growth record reflect the historical or path dependent nature of the growth process. Income divergence results from self-reinforcing vicious and virtuous circles of self-perpetuating slow and fast growth from which it is difficult to escape (although as examples such as South Korea show, not altogether impossible).

BOX 13.3

PATH DEPENDENCE

An economy can be described as path dependent if earlier states of the economy (for example, past rates of growth or unemployment) affect later ones, including what are thought to be long-run states of the economy. In other words, path dependence is a way of saying that 'history matters'. This is far from trivial, because in most economic theories history does not matter, because the commonly used idea of equilibrium usually describes a position to which the economy will return following any disturbance. For example, if the price of a good or service rises (or falls) by any amount above (or below) the price at which supply equals demand, it is usually claimed that it will return automatically to the level that equates supply and demand. In other words, the equilibrium price is ahistorical: it is unaffected by events in the past. According to the principle of path dependence, disequilibrium experienced in the past may affect the value of any equilibrium that the economy later settles into.

Conclusion

Economists disagree as to whether economic growth is a supply-led or demand-led process. Keynesians hold the view that growth is demand-led. Harrodian, Kaleckian and Kaldorian growth theories (named after Roy Harrod, Michał Kalecki and Nicolas Kaldor) all give expression to this view. These theories disagree about the precise source of autonomous demand that is the key determinant or driver of long-term growth. Nevertheless they are all Keynesian and, as such, they all draw similar conclusions about the basic properties of the growth process. These include the paradox of thrift (an increase in the saving rate is harmful to economic growth), and the influence of changes in the distribution of income on growth (including the possibility that raising the share of profits in total income may be harmful to growth).

In Keynesian growth theory, 'demand matters', but 'demand is not all that matters'. Instead, consideration of the causes and consequences of technical progress reveals that, for Keynesians, the interaction of supply and demand is an important determinant of long-run growth outcomes. This interaction may even be crucial to understanding the single most prominent stylized fact that emerges from the historical growth record: namely, the divergence of incomes globally due to forging ahead and falling behind. This prompts one final observation about the growth process. An obvious extension of the

ideas developed and discussed in this chapter is to think about supply-side limits to growth that may arise from the finite nature of the planet and the historical dependence of economic growth on the extraction of exhaustible resources. These themes are explored further in Chapter 20.

Acknowledgements

The author would like to thank the Dana Foundation for financial support that facilitated work on the original version of this chapter. Thanks also to Jesper Jespersen for his invaluable comments on an earlier draft.

REFERENCES

Atkinson, A.B. (2009), 'Factor shares: the principal problem of political economy?', *Oxford Review of Economic Policy*, **25** (1), 3–16.

Cornwall, J. and W. Cornwall (2001), *Capitalist Development in the Twentieth Century: An Evolutionary–Keynesian Analysis*, Cambridge, UK: Cambridge University Press.

Cynamon, B.Z., S.M. Fazzari and M. Setterfield (2013), 'Understanding the Great Recession', in B.Z. Cynamon, S.M. Fazzari and M. Setterfield (eds), *After the Great Recession: The Struggle for Economic Recovery and Growth*, New York: Cambridge University Press, pp. 3–30.

Glyn, A. (2009), 'Functional distribution and inequality', in W. Salverda, B. Nolan and T.M. Smeeding (eds), *Oxford Handbook of Economic Inequality*, Oxford: Oxford University Press, pp. 101–26.

Harvey, J.T. (2014), 'Using the *General Theory* to explain the US business cycle, 1950–2009', *Journal of Post Keynesian Economics*, **36** (3), 391–414.

Kaldor, N. (1985), *Economics Without Equilibrium*, Cardiff: University College of Cardiff Press.

Maddison, A. (1991), *Dynamic Forces in Capitalist Development*, Oxford: Oxford University Press.

Maddison, A. (2008), *Historical Statistics of the World Economy: 1–2008 AD*, available at www.ggdc.net/maddison/Maddison.htm.

Palley, T.I. (2002), 'Economic contradictions coming home to roost? Does the US economy face a long-term aggregate demand generation problem?', *Journal of Post Keynesian Economics*, **25** (1), 9–32.

Palumbo, A. (2009), 'Adjusting theory to reality: the role of aggregate demand in Kaldor's late contributions on economic growth', *Review of Political Economy*, **21** (3), 341–68.

Stern, N. (1991), 'The determinants of growth', *Economic Journal*, **101** (404), 122–33.

Targetti, F. (2000), 'Nicholas Kaldor, 1908–1986', in P. Arestis and M. Sawyer (eds), *A Biographical Dictionary of Dissenting Economists, Second Edition*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 343–52.

Thirlwall, A.P. (1983), 'A plain man's guide to Kaldor's growth laws', *Journal of Post Keynesian Economics*, **5** (3), 345–58.

Verdoorn, P.J. (1949), 'Fattori che regolano lo sviluppo della produttività del lavoro', *L'Industria*, **1**, 3–10.

A PORTRAIT OF NICHOLAS KALDOR (1908–86)

Nicholas Kaldor was born in Budapest, Hungary and became a British citizen in 1934. He joined the faculty of the London School of Economics in 1932, having previously been a student there. In 1950 he moved to the University of Cambridge, United Kingdom (UK), where he became a fellow of King's College (Keynes's old college) before being appointed Professor of Economics in 1966, a position he held until his retirement in 1975. Kaldor was an early convert to Keynes's ideas and made numerous important contributions to Keynesian macroeconomic theory during a distinguished career. For example, he was a progenitor of the idea that the level of the interest rate is anchored to the central bank's policy rate (with the term structure determined by Keynes's theory of liquidity preference), an idea that is central to modern Keynesian monetary theory (see Chapter 4). He was also an active and much sought-after policy advisor on issues ranging from regional development to the design and reform of tax systems (Targetti, 2000). Kaldor advised governments in Asia, Africa and Latin America, but his most prominent position in policy circles was as special advisor to the UK Chancellor of the Exchequer, a role he played twice (in 1964–68 and again in 1974–76) under two different Labour governments.

At the University of Cambridge in the 1950s and 1960s, Kaldor helped to found post-Keynesian or Cambridge macroeconomics, along with other prominent mid-twentieth century Cambridge economists such as Joan Robinson and Piero Sraffa. A particular focus of the early post-Keynesian economists was the extension of Keynes's principle of effective demand to the longer term. Kaldor's work as a member of the post-Keynesian school fell into two distinct phases. During the first (1956–62), he helped to fashion the post-Keynesian theory of economic growth and income distribution, showing how changes in income distribution could reconcile the equilibrium rate of growth with the potential rate of growth in a fully employed economy. Kaldor's theory was an important part of the Keynesian response to early neoclassical growth theory. According to neoclassical theory, the adjustment of the equilibrium rate of growth towards the potential rate of growth results from changes in the capital to output ratio. But according to Kaldor's famous 'stylized facts' of economic growth this is not possible, because the capital to output ratio remains roughly constant in the long term. Kaldor's work on growth theory also contributed to the post-Keynesian theory of the profit rate, a major theme in the so-called 'Cambridge capital controversies' between economists associated with the University of Cambridge in the UK and the Massachusetts Institute of Technology in Cambridge, Massachusetts in the United States. According to the Cambridge, UK economists, the neoclassical theory of the profit rate involved circular reasoning, requiring knowledge of the profit rate in order to calculate the marginal productivity of capital from which the profit rate itself supposedly derived. During this first phase of his contributions to post-Keynesian theory, Kaldor also developed a technical progress function that showed how technical progress could be considered endogenous to the economic growth process. The

➡

relationship between technical progress and economic growth was to become an integral and important feature of Kaldor's later work on economic growth.

During the second phase of his work on post-Keynesian growth theory (1966–86), Kaldor's focus shifted from economic growth and distribution in one country to the determinants of differences in growth rates between countries. His work during this second phase drew on a series of empirical regularities that became known as Kaldor's growth laws (see Thirlwall, 1983). The key ideas that characterized this work (and are reflected in Kaldor's growth laws) include: the importance of export demand as the key driver of aggregate demand formation; the induced effects of economic growth on technical progress (particularly in the manufacturing sector) via Verdoorn's Law; and the positive impact of productivity growth on export demand. These principles combine to produce vicious and virtuous circles of self-reinforcing slow or fast economic growth, which Kaldor used to explain persistent growth rate differences between economies, and the widening gap between rich and poor. The resulting process of 'cumulative causation' was similarly emphasized by the famous Swedish development economist Gunnar Myrdal (who won the Nobel Prize in Economics in 1974), the term itself having originally been used by Thorstein Veblen.

The historical contingency of the economic growth process fed into one of the two major concerns of the final years of Kaldor's life: his critique of conventional equilibrium analysis. Kaldor deplored the stasis of equilibrium analysis, in which outcomes are determined by 'exogenous givens' that, for Kaldor, are not exogenous to the economic system and cannot therefore be taken as given. Ultimately, and in a statement worthy of his Cambridge colleague Joan Robinson, who also wrestled with the antagonism between history and equilibrium, Kaldor declared that 'the only truly exogenous factor is *whatever exists at a given moment of time*, as a heritage of the past' (Kaldor, 1985, p. 61).

The other major theme of the final years of Kaldor's life was his sustained attack on the monetarism of the Chicago School. Inspired by Milton Friedman's rehabilitation of the classical quantity theory of money, monetarism advocated policies designed to curb the growth of the money supply in response to the sudden increase in the rate of inflation experienced throughout the world in the 1970s. According to Kaldor, one of the many errors of this thinking is that the money supply is exogenous (that is, controlled by the central bank), when it is, in fact, endogenous, responding to the demand for and supply of bank credit created by the private sector. Kaldor used his position in the UK House of Lords, to which he had been appointed in 1974, to become a vocal critic of the 'Thatcher experiment', the soon-abandoned attempt by the UK Conservative government led by Prime Minister Margaret Thatcher to pursue monetary policies inspired by monetarism during the late 1970s and early 1980s.

Kaldor is renowned as an economist of great originality who made numerous contributions to economic theory. In addition to those outlined above, he introduced the 'cobweb theorem' to describe the spiral-like process of adjustment towards a market-clearing (supply equals demand) equilibrium that depends on the elasticities of supply and demand. He also made important

contributions to welfare economics that include the compensation principle. This draws attention to the thorny distributional consequences (the creation of 'winners and losers') of even those policy interventions that demonstrably result in net social benefits. At the same time, Kaldor always remained focused on real-world events. He never lost sight of the importance of developing economic theory in the service of doing better applied economics and improving economic policy interventions.

EXAM QUESTIONS

True or false questions

1. According to Keynesian growth theory, only the demand side matters when it comes to explaining economic growth.
2. The historical growth record shows that economic growth as we know it today started well before 1750.
3. In general, global economic growth has involved divergence between the real per capita incomes of the richest and poorest economies.
4. Kaldorian growth theory claims that exports are the key determinant of economic growth.
5. Aggregate demand has no effect on long-run growth in neoclassical growth theory.
6. In Keynesian growth theory, the economy will always be found on its production possibility frontier.
7. Deindustrialization is a common feature of growth once an economy reaches a high level of per capita income.
8. Verdoorn's Law is often understood as an extension of Adam Smith's claim that 'the division of labour depends on the extent of the market'.
9. Keynesians argue that the same source of demand is responsible for driving economic growth throughout history.
10. Harrodian and Kaleckian growth theory share the same ideas about why firms invest.

Multiple choice questions

1. In mainstream neoclassical growth theory, economic growth depends on:
   a) increases in the availability of capital and labour;
   b) increases in the productivity of capital and labour;
   c) both a and b;
   d) neither a nor b.
2. In Keynesian growth theory:
   a) a higher saving rate will lower the rate of economic growth;
   b a higher saving rate will lower the rate of economic growth, but depressing wage income has no effect on economic growth;
   c) raising wage income will lower economic growth;
   d) both a higher saving rate and lower wage income raise economic growth.
3. The historical growth record makes clear that:
   a) economic growth occurs at the same rate everywhere;
   b) economic growth is uneven and also fluctuates between periods;

c) the employment shares of different sectors remain constant as economies grow;
d) by international standards, the richest economies in 1820 grew slowly thereafter.

4. Technical change is important in Keynesian growth theory because:
a) it explains the actual rate of economic growth;
b) it is not important: Keynesians do not pay attention to technical change;
c) it is thought to have stopped after 1820;
d) it links demand and supply in the process of economic growth.

5. The central disagreement among economists about the causes of economic growth concerns:
a) whether growth is determined on the supply side or the demand side of the economy;
b) whether or not technical change matters in the course of growth;
c) the determinants of investment by firms;
d) whether or not international trade affects growth.

6. According to Verdoorn's Law:
a) economic growth has no effect on technical change;
b) technical change is hindered by economic growth;
c) faster economic growth makes for faster growth of labour productivity;
d) none of the above.

7. If economic growth is path dependent:
a) it will not be possible for the economy to grow at an equilibrium rate of growth;
b) economic growth today depends on growth in the past;
c) the supply side has no effect on the growth rate;
d) an economy will grow faster today than it did in the past

8. If income is redistributed towards wages:
a) the rate of growth will slow down according to neoclassical growth theory;
b) there will be no effect on growth in Kaleckian growth theory;
c) the rate of growth may speed up according to Keynesian growth theory;
d) growth theories do not study the relationship between distribution and growth.

9. In Kaldorian growth theory:
a) international trade determines the rate of economic growth;
b) a growing economy will always run a trade surplus;
c) a growing economy will always run a trade deficit;
d) exports are a less important influence on economic growth than investment.

10. 'Catching-up' refers to:
a) the movement towards an equilibrium rate of economic growth;
b) convergence of initially slow-growing economies towards the per capita income level of the richest economies;
c) the process by which economic growth increases over time in Harrodian growth theory;
d) a country that makes a late start to the process of economic growth.

14

Wealth distribution

Omar Hamouda

OVERVIEW

This chapter exposes various economic theories of income distribution: the classical, the neoclassical, and those of Keynes and of Marx. It:

- explains what wealth is and what is understood by wealth distribution;
- points out the distinction between income distribution and income redistribution;
- expands on the microeconomic approach of the neoclassical theory of income distribution based on marginal utilities and productivities;
- argues that the macroeconomic perspective of the neoclassical Keynesian/monetarist model of aggregate demand and aggregate supply provides the justification for factor payments in terms of productivity;
- provides some elementary statistical measures of wealth and income distribution, namely the Lorenz curve and the Gini coefficient;
- applies the Lorenz curve and the Gini coefficient to two cases illustrating relative poverty, that is, the United States and the world;
- presents traditional microeconomic theories of income distribution, notably the classical labour theory of value approach;
- summarizes the contribution of John Maynard Keynes to the rise of macroeconomics and the post-World War II commitment to the welfare state;
- explains Keynes's theory of employment and interest as absolutely essential to his theory of income distribution.

KEYWORDS

- **Classical labour theory of value:** An exchange between two goods or services is made according to the amount of labour used to produce them.
- **Income distribution:** The distribution of the stream of income generated in the production of new wealth, from goods and services, among those factors that have participated in this production (labour, capital and land).
- **Income redistribution:** The reallocation of wealth among members of a community, which might occur regardless of whether the contributors and/or the recipients participate or not in the initial creation of that wealth.
- **Keynesian model:** Aggregate demand (AD) and aggregate supply (AS) correlate the aggregate product Y (or GDP) to the overall price level (P).
- **Liquidity preference (LP):** The choice between holding wealth in cash or lending it with interest; the higher the interest rate, the less is the preference for cash.
- **Marginal efficiency of capital (MEC):** The performance of the productive capital investment in terms of return; this is the concern of the entrepreneur.
- **Marginal propensity to consume (MPC):** The percentage of income received not saved but spent on consumption; this applies to income earners.
- **Neoclassical theory of value:** The desire or utility an individual has for 'things', which determines the goods and services that are exchanged for each other and their relative prices.
- **Wealth:** This constitutes all things desired and valued by a community or an individual and held for direct enjoyment or trade. These things are considered assets, whether in the form of land, capital, supplies, works of art, monies, jewellery, skill, knowledge, information, and so on.

Why are these topics important?

While economists seem to have fewer issues with how resources should be used efficiently to create and increase wealth, they have extremely divergent views about how wealth should be shared. There is an abundance of literature with respect to wealth sharing and about who should decide who should have what. They range from the general to the more detailed, from one perspective to another, from empirically or logically analysed pieces to opinionated discussions. Wealth sharing is at the heart of economics; it is in fact its most important issue, of which its basic principles and its ensuing analysis and policy prescriptions have profound implications for a society's type of socioeconomic organization. Whether it is what share of family inheritance, what factor remuneration in the workplace, which subsidies and social assistance, the ramifications of how such questions are decided shape the basic fabric of any particular society.

Some introductory remarks

Why do police officers earn more than firefighters? Why are two individuals of the same age, in similar circumstances, of similar experience, competence, perhaps of different gender, paid differently? Why is the average annual personal income in Sweden more than US$60 000, while in Zimbabwe it is only US$2000 a year? Why does a large segment of the Nigerian population live on less than US$2 a day, while 1 per cent of the population holds 40 per cent of the nation's wealth? Recently, King Albert of Belgium noted feeling impoverished since his abdication from the throne, because his yearly state sinecure decreased from several million to a mere €900 000. Should he be considered Royally poor? Are older people entitled to a pension whether or not they have contributed to savings?

All these questions have to do with wealth and the rationales for its acquisition and/or redistribution. Through time, thinkers and economists have tried to elucidate all aspects of wealth, and the debates continue. This chapter concerns the distribution of wealth, and as such, it must begin by clarifying what wealth is (Box 14.1): is it given or created, or both? If it is given, why and to whom? If created, how?

Issues related to wealth can be separated into two groups: (1) the distribution of income involves a discussion about its distribution among those involved in the creation of wealth, as well as a discussion about the processes that generate the stream or flow of income to those creating the added value, for their contribution; and (2) the redistribution of income, which is concerned with the allocation of a stock of wealth. Changes in the distribution of the stock of wealth existing at any point in time are handled through legislation, fiscal policy, donation, confiscation and/or coercion. For example, governments may introduce inheritance taxes, tax deductions for gifts to charity or liens on property. Other equally powerful groups may demand extortion money, sell contraband or stolen goods, exact bribes, and so on.

BOX 14.1

INCOME DISTRIBUTION AND INCOME REDISTRIBUTION

It is more accurate to use the term 'income distribution' to refer to how the money generated from the production of goods and services is shared among the factors of production, and the term 'income redistribution' to refer to how monies are levied from some and transferred to others.

Not surprisingly, economic theory has had more to say about income distribution than about its redistribution, as the latter seems to entail a normative stand, in which there are often issues of relative power. Income redistribution is therefore usually associated with fiscal policy (see Chapter 12).

The distribution of a stream or flow of income is traditionally studied under the heading of microeconomics, whose apparatus is used to explain how wages, profits, and rents are determined; while redistribution of the stock of wealth is a component of macroeconomics, involving public finance, social welfare, income transfers, and so on. Since the 1990s there has also been the development of a new subfield of wealth distribution devoted principally to data analysis, touching all aspects of transfers and fiscal arrangements affecting education, pension plans, health, social assistance, gender, child poverty, and so on.

Mainstream economic theory of wealth distribution

For orthodox, mainstream or neoclassical economists, the market economy is the best interactive economic organization, as it results in the most efficient allocation and use of all resources. Income distribution, in a society that consists of a collection of households, is an outcome of a balance of two forces: (1) the sum total of demands, which reflects consumers expressing their needs; and (2) the sum total of supplies, emanating from firms responding to those needs. The core of the model revolves around the households, which, in their choice of distributing their available time between leisure and work, determine how much effort (or employment) they are willing to offer to firms to produce output Y in exchange for a payment of wage w, of which some is used to acquire a share of consumer goods (consumption, $\frac{C}{Y} = cw$) and the rest is saved ($\frac{S}{Y} = (1 - c)w$). c is the marginal propensity to consume, and s is the marginal propensity to save.

The portion that is saved is returned to the firms as investment, a share of which, $\frac{I}{Y} = r$, will yield a return or profit r for future consumption. The income spent on the consumption goods ($\frac{C}{y} = c$) constitutes the returns to the firms to be used for the next period's wages. Since everything produced is either consumed or saved, and what is produced is either consumption goods or investment goods, by definition saving is always equal to investment: $I = S$ (Figure 14.1).

For neoclassical economists, it is the desired needs of consumers that dictate the amount of employment offered, which translates into the employment realized, the production of goods and, finally, the income for satisfying those needs. The income distribution among the factors contributing to the

Figure 14.1 The neoclassical model

production of the final output is determined unambiguously and uniquely by their marginal productivity. The total production of goods and services and the allocated shares of income from that given production are all determined simultaneously.

Hence, we can argue that there is equivalence between:

- what is produced and the income paid to produce goods and services;
- the income paid in the form of money and the total money in the economy;
- the total money and the expenditure of that money on goods and services;
- the expenditure of households and the total receipt of firms;
- total receipts and total output.

The circle is complete, as illustrated by Figure 14.2 (Box 14.2).

Figure 14.2 The circular flow

BOX 14.2

THE CIRCULAR FLOW

Circular flow is a simple description of how goods and services produced are matched by equivalent amounts of payments.

Income distribution and marginal productivity: a microeconomics approach

Most of the foundational premises of neoclassical economics (competition and market forces, the laws of supply and demand and diminishing marginal productivity, and the determination of factor payments) are found in classical economics from Adam Smith to John Stuart Mill, including Ricardo and Marx. The contributions of Carl Menger, Léon Walras and Stanley Jevons shifted emphasis from the labour theory of value (Box 14.3), based on the labour cost of production, to the neoclassical theory of value, based on utility.

The market interaction between supply of and demand for goods and services, considered by the classics to be accidental and not of their concern, became central for neoclassical authors in determining their value theory (Box 14.4).

BOX 14.3

LABOUR THEORY OF VALUE

The labour theory of value is a theory in which the price of a commodity is expressed in terms of its labour cost.

BOX 14.4

THE NEOCLASSICAL THEORY OF VALUE

In the neoclassical theory of value, the price of a commodity is determined by the utility one has for that commodity, and the price of a factor of production is determined by its marginal productivity.

BOX 14.5

MARGINAL UTILITY

Marginal utility is the increased or decreased satisfaction corresponding to the consumption of an additional unit of a good or a service.

Let us suppose two goods, 1 and 2, and their corresponding prices, P_1 and P_2. The ratio is established as:

$$\frac{P_1}{P_2} = \frac{MU_1}{MU_2},$$

which expresses the exchange rate between two goods in terms of their corresponding marginal utility (MU) (Box 14.5). That is to say, exchange between two goods or services is the result of one's desire to acquire something by relinquishing something else of equal utility. Furthermore, the factors of production (labour l and capital k) produce and generate the income used to buy the goods they produce. Individuals choose freely how to divide their available time, according to their desire to work for income or to enjoy unremunerated leisure. It is the offering of their work as service that constitutes the labour supply.

Near the turn of the twentieth century, John Bates Clark applied the law of diminishing marginal productivity to labour and capital to demonstrate that, in neoclassical theory, income distribution among factors of production is uniquely and unambiguously determined (Box 14.6).

Therefore, with the application of this notion to the inputs of production, the quantity of labour and capital used in production is determined by the corresponding marginal productivity of each factor (MP_l, MP_k). If P_k and P_l are the prices of factors k and l, respectively, then:

BOX 14.6

MARGINAL PRODUCTIVITY

Marginal productivity of a factor refers to the increase in production with each additional unit of that factor of production.

Figure 14.3 The neoclassical theory of income distribution

$$\frac{P_l}{P_k} = \frac{MP_l}{MP_k}$$

This ratio indicates that the amount of labour and capital used to produce a given level of output depends on their relative prices, wages (w, the price of labour) and interest (r, the price of capital):

$$\frac{w}{r} = \frac{MP_k}{MP_l}.$$

For a given amount of output produced with a given amount of labour and capital, both w and r are determined by their corresponding productivity. The determination of their prices is illustrated in Figure 14.3.

Whether successive units of labour are applied to a fixed amount of capital, as in Figure 14.3(a), or successive amounts of capital are applied to a fixed level of labour, as in Figure 14.3(b), Clark showed that the payment to labour, determined by MP_l, is given by area *0lbw* (a), while the residual triangle *wbc* represents the payment to capital. When the exercise is repeated, the payment to capital, determined by MP_k, gives area *0kdr* (b), and the residual triangle *rde* represents the payment to labour. The two exercises yield the same results: area *0lbw* determined by MP_l is equivalent to residual area *rde*, and area *0kdr* determined by MP_k is equivalent to residual area *wbc*.

It was also shown that when the marginal productivity approach is applied to determine factor payments, there is 'exhaustion' of the product. In other words, everything produced is distributed among the factors that contributed to the production of that output. In economics, this is referred to as Euler's theorem (Box 14.7).

BOX 14.7

EULER'S THEOREM

Euler's theorem is attributed to the mathematician Leonhard Euler. Used in economics as an 'exhaustion' theorem, it can be stated as follows: total output equals the sum of all the marginal productivities of the inputs times the corresponding amounts of inputs used to produce that output.

BOX 14.8

MARGINAL EFFICIENCY OF CAPITAL

Marginal efficiency of capital (MEC) is Keynes's measure of the performance of productive capital investment, in terms of return.

BOX 14.9

MARGINAL PROPENSITY TO CONSUME

Marginal propensity to consume (MPC) is Keynes's measure, in percentages, of the income received that is not saved but instead spent on consumption.

BOX 14.10

LIQUIDITY PREFERENCE

Liquidity preference (LP) is Keynes's evocation of the choice between holding wealth in cash or lending it with interest.

Both Clark's contribution and Euler's application became the fundamental, powerful premises of factor payment determination in neoclassical microeconomics, found also in today's Keynesian macroeconomics textbooks (Boxes 14.8, 14.9, 14.10, and 14.11).

Although Keynesian economics claims to be derived from John Maynard Keynes's (1936) *General Theory of Employment, Interest and Money*, and

BOX 14.11

NEOCLASSICAL SYNTHESIS

The neoclassical synthesis is a fusion of some of Keynes's insights and neoclassical theory.

indeed uses Keynes's three fundamental concepts – the marginal efficiency of capital (MEC), the marginal propensity to consume (MPC), and liquidity preference (LP) – in its policy analysis it remains nonetheless a macroeconomics offspring of the neoclassical microeconomics model, which became known, as discussed in Chapters 1 and 3, as the neoclassical synthesis. As will be seen below, the Keynesian model and Keynes's own model are fundamentally different.

Keynesians' and monetarists' aggregate demand: a macroeconomics approach

Since the 1960s, Keynesians and monetarists have shared the same neoclassical model in the dominant macroeconomics textbooks (Box 14.12).

There are slight variants reflecting the use of policy applications, and whether these are deemed effective (in the Keynesian perspective) or ineffective (in the monetarist perspective). The model is reduced to a simple relationship of aggregate demand (AD) and aggregate supply (AS) as shown in Figure 14.4.

Both curves, *AD* and *AS*, are independent of one another. *AD* reflects the distribution of the national product in terms of aggregates – consumption, investment, government spending and net exports – while *AS* is derived from the level of employment determined in the labour market, which in turn determines the corresponding level of production of goods and services (through the production function).

BOX 14.12

MONETARISM

Monetarists are a subset of the group of neoclassical economists. They believe in the neutrality of money and the ineffectiveness of introducing monetary policies to alleviate economic business cycles.

Figure 14.4 Aggregate demand and aggregate supply

BOX 14.13

FUNCTIONAL INCOME DISTRIBUTION

The supply and demand relationship of any factor determines the payment to that factor. This is referred to in the literature as functional distribution.

For monetarists, in a world of flexible prices and wages, the intersection of *AD* and *AS* yields output (Y^*) at full employment, Y^f. Keynesians, however, believe that wages, once set, cannot be revised downwards and are hence considered rigid or sticky. This rigidity, a so-called 'market imperfection', results in Y^*, which can be below full employment (Box 14.13). To reduce discrepancy between Y^* and Y^f, Keynesians suggest government policies to maintain high employment or ways of shifting *AD* to the right.

Income distribution in the Keynesian/monetarist model is consistent with the neoclassical microeconomics approach; it does not add much to what is already known, except that the variables are now conceived in the aggregate. Wages are determined in the labour market through the interaction between demand for labour, reflecting the marginal productivity of labour, and supply of labour, expressed in the choice between work and leisure. Determination of the interest rate (in conjunction with liquidity preference, LP) and the rate of return from investment (in relation to MEC) are derived from the

expenditure side, from which *AD* is built. The monetary interest rate and the rate of return on physical investment are indistinguishable. As well, in connection with MPC, which determines how much final product is consumed and how much is saved, is saving (S). Saving has a direct one-to-one relationship with investment (I): $S = I$ always.

The Keynesian macroeconomic theory has little to offer by way of answers to the questions of disparity posed at the beginning of this chapter. Keynesian policies, implemented to achieve certain political targets rather than others, as pursued since the 1960s, are actually at odds with rectifying income disparity, as that trend within and across most nations has been steadily increasing. In sum, successive governments in the industrial economies applying policies in the name of the welfare state, ensuing from neoclassical, monetarist or Keynesian theories, since the 1960s have hardly been successful on the whole in keeping economies out of recession and in providing quality employment (instead creating relatively few well-paid positions, with benefits and security, such as government posts, as opposed to the large number of precarious, temporary and poorly paid jobs in, for example, the agriculture, retail, tourism, and hospitality sectors). This labour opportunity dichotomy is to a large extent responsible for the increasing disparities.

After the immediately ensuing discussion below, Keynes's own theory is suggested as an alternative to the orthodox and other macroeconomic theories. First, however, it is to be noted that theories entailing income distribution have in general stimulated new interest in national accounting and the study of macroeconomic variables, such as national income, aggregate consumption, aggregate investment, and government spending. This has led in parallel to the gathering of more economic data and their analysis.

States of income distribution: a description

In the 1990s and 2000s, studies of all aspects of income distribution have flourished, many focused on the use of data to establish states of income distribution. There are numerous specialized journals that publish this type of study, for example, *Journal of Income Distribution*, *Review of Income and Wealth*, and *Journal of Income Inequality*. Although there are numerous measurement tools to measure inequality, the simplest, primary ones are the Lorenz curve and the Gini coefficient. Each is a specific measure of equality and reflects income distribution within a community, region or country.

Figure 14.5 The Lorenz curve

BOX 14.14

LORENZ CURVE

The Lorenz curve, attributed to Max Lorenz, is a graphic representation relating the percentages of the national income corresponding to groups of population, in percentages of their increasing level of income.

Lorenz curve

One way that economists try to get an idea of how wealth is distributed is to classify the earnings of members of an entity, from their lowest to their highest, by dividing the whole population into percentage groupings (quintiles, quartiles or deciles). Each division is correlated with a corresponding income bracket. This is graphically represented in the Lorenz curve (Figure 14.5), in which the vertical axis represents percentage income, while the horizontal axis represents population distribution in percentages (Box 14.14).

A diagonal line serves to reflect equality: in an ideal world of perfect income distribution, the same percentage of the population receives the same percentage of income. The opposite situation, perfect inequality, would be when one member receives all the entity's income and the rest receives none, as

Figure 14.6 The Gini coefficient

illustrated by the lower horizontal and the right vertical axes of the Lorenz curve.

Curves A, B and C in Figure 14.5 are intermediary cases depicting various degrees of inequality, such that the farther the curve from the diagonal, the higher the degree of inequality.

Gini coefficient

The Gini coefficient (or Gini index) also employs curves in its measurement of income distribution (Figure 14.6 and Box 14.15).

The curve BD describes a given distribution of income farther away or closer to the Lorenz curve diagonal (see Figure 14.5). The shaded area A, defined by the diagonal and by the curve BD, represents a given distribution. Area A becomes larger or smaller, as BD moves closer to the diagonal (perfect equality) or closer to triangle BCD (perfect inequality). The ratio of area A over the area of triangle BCD is defined as the Gini index and serves as a measure of income distribution. This ratio, the Gini coefficient, has a value between zero and one (0 per cent and 100 per cent): it is zero when curve BD is subsumed under the diagonal, and equal to one when it is fused with triangle BCD.

BOX 14.15

GINI COEFFICIENT

The Gini coefficient, attributed to Corrado Gini, is a statistical measure that takes a value between 0 and 1. It indicates a degree of inequality, with higher inequality reflected in a value tending toward 1 and vice versa, when a value tends toward 0.

Figure 14.7 A sample of Gini coefficients

Just like the Lorenz curve, the Gini index can be used to describe income distribution within a single entity but also to compare disparities among entities. Figure 14.7 gives the Gini coefficient for a sample of countries. Norway has the lowest coefficient, which means that it has a more egalitarian income distribution. South Africa, on the other hand, has a higher number, reflecting a rather less egalitarian distribution. Gini coefficients are calculated for a given period. Figure 14.7 compares coefficients for two distinct periods, 2006 and 2011, showing deterioration in the disparity of income in almost every country noted, save South Africa and Norway, whose numbers did not change. In different periods in recent history, one can find country cases when there has been either deterioration followed by improvement or the reverse.

As an illustration of income disparity using the Lorenz curve for a particular country, Figure 14.8 shows data for the United States. Income levels in the

Source: Tax Policy Center (the dataset refers to May 2011; available at http://taxpolicycenter.org/numbers/displayatab.cfm?DocID=2970).

Figure 14.8 The Lorenz curve for the United States

United States are depicted on the vertical axis; on the horizontal axis are the population groupings.

If one takes the gross national income (GNI) in 2011, which has been generated by that year's total United States gross domestic product (GDP), in goods and services, and divides the GNI by the total population, one obtains the average per capita (per person) revenue in the United States. Figure 14.8 shows that average earnings (the 50 per cent mark) were about US$42 000 in 2011. It is obvious that this average earning is just an indicator, as is the Gini index for the United States (at about 47 per cent; see Figure 14.7), both being ways of measuring income disparity in a nation. A closer look at the data in Figure 14.8 reveals, however, a more detailed understanding of the disparity in earnings among Americans in 2011. The curve shows that 20 per cent of the poorest Americans earned less than US$17 000 that year, while 10 per cent of the richest earned more than US$360 000 in the same period.

A Lorenz curve can also be constructed for the distribution of world income. In Figure 14.9 one can see that the income distribution curve for the world as a whole is much more pronounced than that of the United States and describes even greater inequality. It is interesting to notice that all those in

Figure 14.9 The Lorenz curve for the world

the United States earning more than US$10 000 and less than US$20 000, considered to be part of the lowest 30 per cent poor in Figure 14.8, find themselves to be part of the 20 per cent richest group, in the perspective of world income. More striking is that all those earning more than US$40 000 in the United States (which is also about the average income per capita in many developed countries) are among the 10 per cent richest people in the world, compared to 30 per cent of the world's population (2.3 billion people), who live on less than US$2 a day. The two curves together illustrate that there are different understandings of what it means to be 'poor' or 'rich', depending on whether one draws conclusions about disparities through national or international data. Being rich or poor in different parts of the world does not have the same significance worldwide.

In just what way the Lorenz curve, the Gini coefficient, and other broad indices such as the Index of Human Development (IHD), describe world income distribution (what it is, what it means and how disparities have come about) is the subject of continuing study and debate in many areas of economics, such as economic development, production and economic growth, and industrialization and productivity. More in-depth empirical evidence shows that disparities in income and wealth distribution are diverse and changing through time in various directions. Theorizing these disparities requires models whose main variables reflect the interaction between income earners, as contributors to the creation of wealth and as recipients of a share of that wealth. Understanding such models begins here with the

return to classical economics, whose primary concern was the discovery of the laws governing factor payments.

Heterodox perspectives

Despite the existence of several alternative heterodox theories of wealth distribution, almost all use the basic framework and premises of the economics of either the classics or Keynes.

Figure 14.10 The classical/Keynes model

The classics' and Keynes's models are constructed as a tripartite balance of forces of supply and demand for factors of production, based on the interactions among rentiers, entrepreneurs, and workers. Rentiers provide land (in the classics) or financing (in Keynes) in exchange for a return (R), entrepreneurs lead and manage production and expect a profit (r), and workers offer labour for a payment of wage (w) for their productive effort (Box 14.16). All three contribute to the creation of a certain level of product Y and receive a share $\frac{R}{Y}$, $\frac{r}{Y}$, and $\frac{w}{Y}$ of the total product: the share of rent, of profit, and of wages, respectively; all shares add to 1 (Figure 14.10).

The classics studied income distribution from a microeconomics perspective. The ongoing production of goods and services generates a flow of income. In any specific period of time there is a given level of income specific

BOX 14.16

NATIONAL ACCOUNTING

The data for the total value of these shares – wages (w), profits (r), and rents (R) – are usually found in the national accounting tables as gross national income (GNI), as the counterpart of the gross national product (GNP) or the gross domestic product (GDP).

to that level of production, which is divided among the factors of production that contributed to this outcome.

Classical economics and the law governing factor payment

While classical economics was discussed in Chapter 3, it is worth highlighting it here for the purposes at hand. Assume a farmer who rents land and with the help of one worker produces 1 bushel of grain valued at P, its price or its return. That value is shared between the worker, who provided his labour (L) (at wage, w), the landowner or rentier, who supplied the land (for rent, R), and the farmer or entrepreneur, who invests capital (K) (at profit r). Therefore, in the search to discover the laws that determine the distribution of income among the various factors of production:

$$P = w + r + R$$

What is the share of each party? For Adam Smith the use of land presents two characteristics: (1) not all land has the same yield, thus two parcels of the same size may generate two different rents (R_1 and R_2), with rents set depending on the fertility of the land; and (2) following the law of decreasing marginal productivity, applying more and more amounts of labour and capital to the same parcel of land results in less and less marginal product. In separating which returns go to capital and which to labour, the classics assumed wages (w) to be given at their minimum; namely, at the subsistence level. Figure 14.11 illustrates how the distribution of income to each factor is determined, in the different contexts of differing qualities of land.

Classical economists made three important assumptions: (1) when labour is abundant, as it was in their time, wages (w) will tend to become uniformly

Figure 14.11 The classical distribution of income

distributed albeit at their minimum; (2) competition for work produces a uniform level of wages (w^*), which is tantamount to a horizontal labour supply (w^*w^*); and (3) since capital flows where the return (r) is highest, competition for capital creates forces that tend to equalize profits (r^*). Suppose now parcels of land of different yields, A, B, . . ., n, to which amounts of labour (L) and capital (K) are applied. As shown in Figure 14.11, wages (w^*) and profits (r^*) are the same for the use of land A, B or n. In land A, if $L_A + K_A$ is utilized, land A will generate an amount of total product equal to area *0cde*. If $0w^*$ is the wage rate, and w^*r^* is the profit rate, then area $0cw^*w^*$ represents the wage bill, $w^*w^*dr^*$ is the total profit, and the residual area der^* is rent.

It is clear from Figure 14.11 that wealth (the sum of the total product for each parcel of land) is considered to be created from a given endowment (of labour, capital and finance), and that the product is distributed entirely among its three factors: workers, entrepreneurs (or capitalists) and rentiers (or financiers). As more and more factors are used on the same area of land, productivity decreases. The thick black line of each diagram, which is lower for land B than for land A, and for land n than for land B, represents the diminishing or marginal productivity (MP) of labour and capital in the utilization of each land. As for rent, it is higher for land A than for land B. In the case of the marginal land n, the first units of labour and capital produce just enough to pay for wages and profits, but they do not generate enough for rent. Thus, the owners of lands of type n would have to rent them for free or not at all.

According to the classics, competition among factors of production dictates that it is the last dose of combined capital and labour applied which produces a given yield. From Figure 14.11 one also sees that if rent is considered residual, as the classical economists assumed it to be, then it is the marginal land that determines the marginal product of capital and labour. Hence:

$$P = w + r$$

or more appropriately, if it takes L labour and K capital to produce a certain amount of grain, then the expression can be rewritten as:

$$P = wL + rK$$

Classical economists proceeded to a simplification of this expression by assuming that capital K is produced with both labour and capital at time t–1, and K_{t-1} is produced with labour and capital at time t–2, and so on. They

thereby reduced K to its labour content and called it indirect labour. Hence P can be written as:

$$P = wL^*$$

where L^* is the sum of direct and indirect labour (see Chapter 3).

Suppose two goods, 1 and 2, and their corresponding prices, P_1 and P_2. In the classical labour theory of value, the ratio:

$$\frac{P_1}{P_2} = \frac{wL^*_1}{wL^*_2} = \frac{L^*_1}{L^*_2}$$

expresses the exchange rate of two goods in terms of the labour (direct and indirect) used to produce them.

For classical economists, from Adam Smith to John Stuart Mill, including Karl Marx (see the portrait of Marx at the end of this chapter), income distribution is tied to the study of production and the contribution of each class. The logical analysis entailed in the labour theory of value resulted in controversial debates about the source of the creation of wealth and the justification of the returns to each class. The controversy among schools of classical thought centred on numerous questions. Which among the three factors (labour, capital and land) is the creator of wealth? If the wage bill is fixed, can wage pressure be relieved only by satisfying the demands of one group of workers at the expense of another? And, more generally, should private property be allowed? Violent, diverging views were held, which consequently had a profound influence on ideas about the merits of the market economy versus collectivism. For Marx, for example, for whom labour is the sole creator of value, the entire return from production ought to go to labour, and workers should have full control of the means of production.

Pre-World War II economics was polarized. On the one side, dominant neoclassical economics, in the tradition of its classical predecessor, relied mainly on market mechanisms to justify income distribution. On the other side was the Marxian perspective, which rejected the validity of deterministic market forces and its consequent distribution. The severity of the crisis of 1929 and the revealed inability of economic theory to help find ways of getting an economy out of depression, led some economists, in particular John Maynard Keynes, to think of alternative ideas.

BOX 14.17

EFFECTIVE DEMAND

Keynes's notion of 'effective demand' corresponds to the meeting point of 'aggregate demand' and 'aggregate supply price', a point where both the demand side and the supply side are as perceived and assessed by the entrepreneur.

A new macroeconomics approach: stock of wealth and social well-being

The publication of *The General Theory of Employment, Interest, and Money* by John Maynard Keynes in 1936 radically changed the study of economics, in both its theory and its policy. Keynes provided a new approach based on macroeconomic variables and on social accounting. Rather than relying on individual supply and demand, as was the case in the previous study of economics, Keynes shifted the emphasis to aggregate supply and aggregate demand. Keynes's general idea, simple in its essence but difficult and cumbersome in its conception, created much confusion and a rift between his theory and the corrupting Keynesian interpretation.

Keynes called 'effective demand' the productive output brought about through the interactions among rentiers who provide financing, entrepreneurs who manage production, and workers who offer labour (Box 14.17).

In a monetary economy of production, aggregate employment and income depend mostly on the entrepreneurs, who anticipate what they think are the needs of the rest of the community and then invest in the employment and the production from which national wealth is created. The entrepreneurs' anticipated income may or may not generate production that satisfies all desired needs, and/or may satisfy the needs of some more than others. In Keynes's theory, labour costs and production precede income; investments are thus made on the basis of expected future income. Entrepreneurs' decisions to produce or not and with how much labour depend mostly on anticipated profits.

For Keynes, it is the entrepreneurs' assessment of effective demand that determines employment and income distribution. Keynes's aggregate demand and aggregate supply correlate a level of employment to expected proceeds or return on investment. Labour is passive in determining investment. Finance, however vital, is in constant flux, and the cause of instability

in the economy is the volatility of investment. It is caused by the attitude toward liquidity determined by financial activity.

Money, credit contraction and expansion, and financial speculation are the major variables affecting the determination of income distribution, but they also cause the redistribution of income among wages, profits, and rents. Employment depends on investment, which in turns depends on the availability of finance. Money, or liquidity, can be used as a tool in the investment process to achieve both employment and equitable distribution of income among the three classes – labour, entrepreneurs, and financiers – who create value-added. Inspired by the classics, who built their model of distribution on the hypothesis of the scarcity of land – understandable for their time since theirs was an agrarian economy – Keynes came to realize that the economy of the early twentieth century was industrial and based on a sophisticated financial system. The alternance of abundance and scarcity of liquidity was the major cause for its violent fluctuations in both production and employment. Keynes in a clever manner used the classical model and transformed it by making the financial liquidity required for investment the main source of the economy's creating growth and employment.

Keynes provided a theoretical mechanism that allows for the creation of employment and income in a monetary market economy. For Keynes, aggregate profit, more as means than end, is his theoretical starting point. As in Figure 14.12, employment is thus correlated to 'expected proceeds'; namely, the return on investment.

Figure 14.12 Keynes's effective demand

Figure 14.13 The marginal efficiency of capital

Keynes, in an original way, applied thus the classics' marginal productivity to investment, measured in money, for a given production and employment. The amount of investment determines returns on investment; wages are residual. Just as in the classical case with different land (see above), here there are different production–employment processes to which a level of investment is applied.

Keynes constructed a prospective yield schedule or curve of the corresponding diminishing productivity of various investments in the aggregate (Figure 14.13). He contrasted that schedule to the supply price curve, which describes how returns are divided and paid back to the rentiers (and entrepreneurs) as interest, dividends, and costs related to the utilization of capital (replacement, amortization, administration, and so on).

In Figure 14.13, use of investment I^*, which corresponds to a given level of employment, generates production income area $0CI^*E_0$, which is divided between investment (area BOI^*E_0) and wages (area CBE_0). To increase the level of employment, investment must increase. One way is to shift the

prospective yield curve to the right, which can be achieved with an increase in productivity, an increase in prices or a decrease in taxes. A second way would be as a result of shifts in the supply price curve to the right, as a consequence of a lower cost for borrowing funds or replacing equipment. Both these shifts permit movement of the intersection E_0 to E_1'. With more investment and more employment, the triangle CBE_0 (the total wage bill) increases, while area BOI^*E_0 (the payment to rentiers) becomes smaller. The second way to increase investment is through a policy of lowering the interest rate.

Keynes's most novel proposal was in the area of monetary policy. He felt that money, being legal tender ('state money'), should be made available to entrepreneurs at low interest rates to encourage investment, for them to be able to maintain a marginal efficiency of capital (MEC) or return on their investment high enough to stimulate more investment. An abundance of available money would counterbalance financial lenders, who would otherwise take advantage of tight money or liquidity preference (LP) to extract high interest. It is in this context that Keynes (1936) introduced his famous expression 'the euthanasia of the rentier', meaning that the monetary authorities would thereby keep the financial sector in check. Maintaining a high level of investment is a prerequisite to increasing employment and improving wages. As long as the economy is near (but not at) full employment, a hand-in-hand fiscal policy directed toward the community's (marginal) propensity to consume (MPC) would keep inflation in check.

Central to Keynes's theory of income distribution are the concepts of MEC, LP, and MPC (see above, and the keywords for this chapter). They are built into his prospective yield schedule and supply price curve, as well as the interrelated aggregate demand and aggregate supply that constitute effective demand. All three are perceived to impact firms' investment and employment. The MEC, LP, and MPC in the Keynesians' and the monetarists' models filter their impact through AD, which is considered to be independent of AS. For these models, as part of AD, they boost any expenditure forming GDP, of which consumption is the largest component.

In conclusion, it is obvious that the classical/Keynes model of a market balance of forces is more appropriate than that of the orthodox model to address the general aspects of income distribution among the three social classes (Figure 14.14).

The answers to the questions posed at the beginning cannot be found in economics alone, but require consideration of the political institutional setting, which defines the relative powers impacting upon the distribution

Figure 14.14 Relative economic power groups

and redistribution of income and wealth. In a monetary economy, the legal management of state money rests with financial institutions, which gives banks considerable privileges in the creation and channelling of credit, not always in the best interests of all. For his policy to be effective, Keynes felt that the banking system must be separate and independent of the financial sector. Although private, it would remain nonetheless an arm of a nation's central bank; its unique role, however, would simply be to provide liquidity to the economy.

Without the legislation permitting limits on corporate liability and the amassing of huge capacity by financiers, capital accumulation and takeovers would be difficult. Without conferring concession rights in the workplace, unions would have a hard time protecting the interests of their particular members. Present situations, such as the emergence of powerful poles of decision-making (the 'monoavari' in finance, monopolies in industry, and the 'monomani' in the labour market); the relegating of the risk factor to government; the conferring of concessions, licences, and benefits to particular groups; all contribute to hierarchies of relative power within a society, which explain the variety of disparities.

Some concluding remarks

The establishment of the welfare state in the 1950s and 1960s, with sustained economic growth and high levels of employment as its main economic goals, has meant a continuous commitment to specific Keynesian policies. Policy guidance from Keynesian economic theory in order to stimulate aggregate demand, consisting of adopting measures affecting distribution through

MEC, MPC, and LP, has resulted in increased government spending and the expansion of credit for mass consumption to be used as the engine of economic growth since the 1980s. Consumption, to the point of addiction, encouraged and developed by the rentiers, has shifted the distribution of income from both wage earners and entrepreneurs to their favour. There is growing evidence that Keynesian policies are increasing income disparity, even though they temporarily dampen the effects of recessions. Further, new practices of indebtedness, which consist in abetting consumption first, before the production and generation of income, with all the uncertainty this entails, have drastically changed attitudes toward effort and production, promoting short-term gain and the 'quick exit' rather than long-term stability.

Keynes's unorthodox theory offers an alternative: encouraging policy measures, affecting MEC, MPC, and LP, in order to stimulate effective demand from the production perspective. That is to say, production, the necessary securing of employment, and the ensuing generation of income come first, before consumption. State money made available to entrepreneurs and an interest rate held low, along with a profitable rate of MEC, are the conditions that encourage investment and in turn employment. They will occur only if the functions of the banking system are separate from those of speculative financial institutions and under the strict control of the central bank to manage the legal tender for the good of society as a whole.

It is projected that by 2050, in an ever more digitized world, 70 per cent of the world's population of 9 billion will be living in megacities, thus increasing the interdependence of people both for their demand for employment and income, and for their procurement of food, clothing, and shelter. The scale of the world aggregate demand and aggregate supply puts tremendous pressure on global resources and encourages capital accumulation, which develops more dependency on finance. Given economic circumstances, the banking system and the financial sector find themselves in an absolute position to direct investment. Unchecked, as Keynes predicted, they will continue to extract high rent.

It seems that the wonders of technological change are multiple: they allow for labour saving, they liberate the workforce from hard employment and offer it more leisure time, and they create affordable mass production, and together with a favourable financial environment offering easy credit, foster consumption. At the same time, this reflects a chain reaction in which finance receives high returns, the saving of labour keeps a lid on wages, and industry, through rationalization, attempts to prevent its MEC from declining.

The well-being of a society, in terms of both economic and social stability, rests to a large extent on an equitable distribution of income. Keynes reflects a balance between the various members of society (rentiers, entrepreneurs, and workers), whose means and ends permit their participation in the creation of wealth. The key to reach that goal was for Keynes neither speculation nor avid consumption, but employment, the exertion of effort in production.

REFERENCES

Keynes, J.M. (1936), *The General Theory of Employment, Interest, and Money*, London: Macmillan.
Marx, K. (1867), *Das Kapital, Volume I*, Hamburg: O. Meissner.
Marx, K. (1885), *Das Kapital, Volume II*, Hamburg: O. Meissner.
Marx, K. (1894), *Das Kapital, Volume III*, Hamburg: O. Meissner.
Marx, K. and F. Engels (1848), *Manifest der Kommunistischen Partei*, London: J.E. Burghard.

A PORTRAIT OF KARL MARX (1818–83)

The life of Karl Marx, a German philosopher and economist, spanned the nineteenth century, from the rise of the industrial revolution to the apogee of free market capitalism. Marx was appalled by the harsh impact of the progress of industrialization on the lives of ordinary people. Marx's ideas of worker exploitation, mass pauperization, and an increasing gap between the accumulated wealth of a few and the misery of the large number, had a profound result for economics. Critical of individualism, private property, the laissez-faire of the Physiocrats, and then the guiding classical economics of Adam Smith and David Ricardo, Marx advocated collectivism, cooperation in production, and communal ownership, which is to be understood as communism. Although most of his thinking was devoted to criticism of capitalism, he dwelt little on theoretical alternatives to reform the free market economy.

Marxism, after Marx's name, became synonymous with class struggle and full control of the means of production as its ultimate goal. Antithetical to the liberal free market ideas of Adam Smith and orthodox economics, Marxism created, with its rise, an irreconcilable clash between two visions of economic and sociopolitical organization to promote the betterment of human beings. In the twentieth century, attempts to apply Marxism began with the Russian Bolshevik Revolution in 1917, which led to the establishment of state socialism. The post-World War II polarization of the world and the subsequent Cold War dividing East from West were the direct consequence of these antagonistic visions.

Marx's contributions came in two waves. In a first, early phase, he believed in the possibility of changing the course of capitalism. With Fredrick Engels, he participated in drafting the famous *Manifesto of the Communist Party* (Marx and Engels, 1848). Trained as an editor, Marx used his legal and journalistic skills to promote his ideas. Active in organizing the working class to rise up and take control of their destiny, he fomented the Revolution of 1848, which resulted in the bloodshed and repression that led to his exile from Germany. In a second phase, having retreated to London, Marx devoted himself to writing. He produced his *magnum opus*, the three-volume *Das Kapital*, a theoretical explanation and justification for the rejection of capitalism, which, he believed, contains 'the seeds of its own destruction' (Marx, 1867, 1885, 1894).

Marx had a specific take on income distribution. It is explained very briefly in the text above how Marx thought of labour as the sole creator of value. The entire net product, after the cost of the means of production are subtracted, should belong to the workers who contributed to that production. Marx's theory of income distribution rests on two hypotheses: (1) the transformation of values into prices; and (2) the long-term diminishing rate of profit. Marx used these two tenets to criticize capitalism, to the point of predicting that sooner or later it would collapse. When that happens, workers will take control of the means of production and the return from whatever they produce will be theirs.

In the context of capitalism, with regard to hypothesis (1), at the cost of production level, workers get only one share of the value of the product they produce. The remaining goes to the payment of the

➡

entrepreneurs and financiers, the owners of capital. That same product, when brought to the market by an intermediary, who had previously acquired it at its wholesale price, gets resold at a higher, retail price. For Marx, the workers, the sole creators of value, are thus exploited twice: at the production level, where they get only a fraction of the value of the product they created, and then in the marketplace, where they buy a fraction of that product at the retail price, which results in their paying more (than what they earned for that fraction) as the purchasing power of their income is further diminished.

As for hypothesis (2), the definition of the rate of profit π is given by the difference between the total earning, E, minus total cost, C, per dollar invested:

$$\pi = (E\text{-}C)/C$$

C is made of the wage bill (wL) and capital expenses (rK).

Classical economists came to the conclusion that the rate of profit in each industry is falling over time, due to competition among capitalists to acquire an ever bigger share of the market. The difference ($E - C$) is the surplus, S, or profit, and is what allows firms to innovate and increase productivity. Marx had a different interpretation of that same formula. If both the numerator and denominator of π are divided by wL, then the numerator becomes the surplus per worker, which Marx terms the rate of exploitation, and the denominator, the capital/labour ratio, which he calls the organic composition of capital. Over time, there is a limit to how much can be squeezed out of labour to create a surplus. Capital, however, tends to increase much faster. This makes the growth of the denominator of π increase more than that of the numerator, thus diminishing the rate of profit. The only way for capitalist firms to survive is therefore to keep wages at their minimum, to make up for falling profits.

EXAM QUESTIONS

True or false questions

1. In the last decades, since the advent of the cyber-technology age, income disparity has been declining due to the democratization of knowledge.
2. There is no distinction between the average and the marginal productivity of a factor when it comes to measuring income disparity.
3. In neoclassical economics, the exhaustion theorem asserts that all the factors are justly paid according to their marginal productivity contribution.
4. For neoclassical economists, it is the desired needs of consumers that dictate the amount of employment on offer.
5. Monetarists are a subset group of neoclassical economists who do not believe in the neutrality of money and are therefore strong supporters of banking regulation.
6. In Keynes's theory, the entrepreneurs' motivation for profits is not greed but rather the obtaining of the means to increase employment and, by extension, more equitable income distribution.
7. In classical economics, when government regulates the markets, they in turn regulate income distribution: through government policy, wages and profits are uniformly distributed, and capital flows whither the return is highest.

8. The major shift that Keynes's methodology introduced was for economic theory to move from consideration of the macro to consideration of the micro aspects of an economy.
9. According to Marx, once the basic costs for capital are paid, all income should be distributed among the workers.
10. Keynes conceived the marginal efficiency of capital as the rate of return that makes it just worthwhile to pay for one more unit of investment capital.

Multiple choice questions

1. A free market economy allows for:
   a) free mobility of labour;
   b) free flow of capital;
   c) money to serve as the sole medium of exchange;
   d) all the above.
2. The Lorenz curve:
   a) Is an apparatus that helps to understand why there is income disparity among nations.
   b) Is a formula, which compares different percentages of the population to their respective incomes, from the poorest to the richest.
   c) Is a graphical relationship, which associates different percentages of groups of the population to their respective incomes as a percentage of the total population income.
   d) Permits calculations of income per capita for different income groups.
3. The Gini coefficient is:
   a) A formula used to calculate what is needed to diminish the disparity of income between the poor and the rich.
   b) A ratio of the expected income inequality over the actual inequality.
   c) An income disparity index, which takes a value between one and zero, such that as the value approaches zero, it indicates greater equality.
   d) An income disparity index, which takes a value between one and zero, such that as the value approaches zero, it indicates greater inequality.
4. A functional distribution stipulates that workers' wages be:
   a) The amount of the relative money paid to workers depending on their occupation.
   b) The amount of money paid to workers to stimulate them to work.
   c) Determined by supply and demand in the labour market.
   d) Higher as their productivity diminishes, to compensate for the drop.
5. In the orthodox theory of value, prices are determined according to:
   a) The average satisfaction or utility.
   b) The marginal satisfaction or utility.
   c) The total satisfaction or utility.
   d) The quantity of labour it takes to produce a good or service.
6. According to the classics, the long-term rate of profit for each industry:
   a) Is not predictable due to market volatility.
   b) Stays constant due to competition among the large number of firms.
   c) Tends to diminish due to competition and diminishing marginal productivity.
   d) Tends to increase due to diminishing costs over time.
7. For Keynesians the effective demand:
   a) Corresponds to the output level and corresponding general price level where aggregate demand meets aggregate supply.

b) Corresponds to the employment level and the corresponding prospective proceeds where aggregate demand meets aggregate supply.
c) Reflects the fulfilment of the aggregate demand as an expression of consumer satisfaction independently of aggregate supply.
d) Reflects the fulfilment of the aggregate supply as an expression of producer satisfaction independently of aggregate demand.

8. Keynes's effective demand:
a) Corresponds to the output level and corresponding general price level where aggregate demand meets aggregate supply.
b) Corresponds to the employment level and the corresponding prospective proceeds where aggregate demand meets aggregate supply.
c) Reflects the fulfilment of the aggregate demand as an expression of consumer satisfaction independently of aggregate supply.
d) Reflects the fulfilment of the aggregate supply as an expression of producer satisfaction independently of aggregate demand.

9. Marx's concept of the exploitation rate is:
a) The rate at which resources are depleted.
b) The rate at which technology is introduced by the capitalist.
c) The speed at which a worker produces.
d) The surplus value that a capitalist can extract from each worker.

10. To improve on income distribution, Keynes proposed:
a) To drive the monetary interest rate down by making more liquidity available to entrepreneurs.
b) To increase the monetary interest rate so that liquidity can be distributed more efficiently in the financial market.
c) To increase the monetary interest rate to encourage savings and then, in turn, investment and employment.
d) To let the banking institutions determine the best rate of monetary interest, as they have a better idea of what is most appropriate.

Part IV

International economy

15

International trade and development

Robert A. Blecker

OVERVIEW

This chapter:

- explains the conventional theory of comparative advantage, which implies that free trade benefits all countries, but also shows that trade creates losers as well as winners and may increase inequality within countries;
- discusses how that theory rests on the unrealistic assumptions of balanced trade and full employment, without which the theoretical 'gains from trade' may not be realized;
- shows that changes in the terms of trade (international prices) redistribute the gains from trade between nations, for example, when increases in primary commodity prices benefit commodity exporters at the expense of commodity importers (or the opposite when those prices fall);
- presents the heterodox alternative, in which trade often follows absolute rather than comparative advantages; when this occurs, countries that achieve more rapid export growth and/or trade surpluses may benefit at the expense of other countries;
- examines the reasons why countries sometimes impose trade barriers such as tariffs (taxes on imports), especially to protect 'infant industries' in the development process, while on other occasions nations cooperate to sign trade agreements that reduce or eliminate those barriers;
- discusses the origins of the 'new economic nationalism' that led to Brexit (the United Kingdom leaving the European Union) and former United States (US) President Donald Trump's trade war with China and other 'America first' policies.

KEYWORDS

- **Absolute competitive advantage:** A country having the lowest monetary cost of production of a good, as a result of low wages, a low currency value and/or high productivity of labour.
- **Comparative advantage:** A country having the lowest relative (opportunity) cost of producing a good, meaning that it has to give up less of other goods to produce it than other countries.
- **Global value chain:** The sequencing of productive activities for a firm, from research and development (R&D) and product design, through sourcing of raw materials and other inputs, production of intermediate goods (parts and components), assembly of final goods, packaging, transportation, distribution, sales and after-market service, all of which may be carried out in different countries.
- **Infant-industry protection:** Using trade barriers such as tariffs (taxes on imports) and quotas (quantitative limits on imports) to promote the development of new industries that can eventually become internationally competitive.
- **Tariffs:** Taxes on imports, which make them more expensive for domestic consumers to purchase and make it more profitable to produce the goods domestically.
- **Terms of trade:** The relative proportion in which goods are exchanged internationally (the amount of imports a country can buy with its exports); especially important for exporters of primary commodities (agricultural and mineral products).
- **Trade balance:** The difference between the value of a country's exports and imports of goods and services; a positive balance is called a surplus and a negative balance is a deficit.
- **Trade liberalization:** Reducing trade barriers such as tariffs (import taxes) and quotas (quantitative restrictions), either through multilateral negotiations or via regional/preferential trade agreements.
- **Trade war:** A situation in which countries impose tariffs and other trade barriers on each other in retaliation for the other's tariffs or actions, leading to a reduction in the volume of global trade.
- **Value added:** The extra value added to inputs (raw materials and intermediate goods) by labour and capital in production; it equals the difference between the sales price and non-labour variable costs.

Why are these topics important?

International trade has been an important feature of the economic growth process since the dawn of modern capitalism. The colonial empires of the sixteenth to the early twentieth centuries were built on a very unequal form of trade, primarily involving the exchange of natural resources from the colonial areas for manufactured goods from the imperial powers. The countries

that have grown most rapidly since the mid-twentieth century, such as Japan, South Korea and China, have relied on exports of manufactures as a key part of their growth strategy. Supporters of the global trading system argue that unfettered trade leads countries to specialize in the goods and services that they can produce most efficiently, resulting in higher productivity and increased consumption levels in all countries. Critics, however, charge that the global trading system often widens income gaps between more advanced and less developed nations, and can also exacerbate inequality in the distribution of income within countries (see Chapter 14).

The share of internationally traded goods and services in global gross domestic product (GDP) increased dramatically between the 1960s and early 2000s, before levelling off after the global financial crisis of 2007–08, as shown in Figure 15.1. This tremendous expansion of trade was driven by several factors, including deliberate efforts by governments to open up markets by negotiating reductions in trade barriers – a process known as trade liberalization – as well as changes in technology that have revolutionized how products are made, how cheaply goods can be shipped, and how easily firms can communicate with production facilities around the world.[1]

Source: World Bank (2019).

Figure 15.1 World trade in goods and services as a percentage of world GDP, 1960–2017

This chapter is concerned with theories that try to identify the relationship of international trade to long-run economic growth objectives and other core macroeconomic issues such as unemployment.

The orthodox approach: the theory of comparative advantage

The orthodox approach to international trade is known as the theory of comparative advantage. This section will explain this theory and some of its limitations, while later sections in this chapter will explore heterodox alternatives. The essential idea of comparative advantage is that every country should specialize in (and export) the goods that it can produce with the relatively lowest cost compared to other countries, while importing those goods that can be produced at a relatively lower cost abroad. The key question then is how to define the meaning of goods being relatively cheaper or more expensive. Based on the classical statement of the theory of comparative advantage by David Ricardo (1821/1951), a simple and compelling way to define this concept is in terms of relative labour cost, that is, by comparing the cost in labour time of producing goods in each country, as in the following example.

An example of comparative advantage

To illustrate this theory, consider a simple model (intended to represent the world economy around the 1950s) of two countries, the United States of America (USA) and East Asia (EA), which can produce two goods (televisions and rice) with the labour costs (person hours per unit of output) shown in Table 15.1. Given these (purely hypothetical) numbers, the relative cost of producing a television in the USA is only 2/4 = 0.5, which is the ratio of the hours of labour required to produce one television (2 hours) to the hours required to produce 1 ton of rice (4 hours). This ratio is very important, as it corresponds to the microeconomic concept of opportunity cost: the USA has to give up 0.5 tons of rice in order to produce each additional television.

Table 15.1 Example of Ricardian comparative advantage

	Hours of labour required to produce:	
	Televisions (per unit)	Rice (per ton)
United States	2	4
East Asia	6	5

In EA, however, the relative (opportunity) cost of a television is 6/5 = 1.2 tons of rice (because it takes 6 labour hours to produce a television and 5 labour hours to produce 1 ton of rice), so televisions are relatively more expensive to produce there (EA has to give up 1.2 tons of rice for each television it produces, compared with only 0.5 in the USA). Hence, we say that the USA has the comparative advantage in televisions. Because this is a purely relative comparison, the same logic implies that EA has the comparative advantage in rice. To see this, note that the opportunity cost of producing rice in the USA is 4/2 = 2.0 (the reciprocal of the US opportunity cost for televisions), while the opportunity cost of rice in EA is 5/6 = 0.83 (the reciprocal of the EA opportunity cost for televisions). Since 0.83 is less than 2.0, we can see that EA has the relatively lower opportunity cost for rice and hence will export it.

Note that, according to the theory of comparative advantage, a country does not need to have an absolute productivity advantage (that is, the lowest labour time per unit, or highest output per hour) in a good in order to export it. In our example, EA exports rice even though it has an absolute disadvantage in rice (it takes 5 hours of labour to grow 1 ton of rice in EA, compared with only 4 hours in the USA). Indeed, in our example the USA has the absolute advantage in both goods, but it has a comparative advantage only in televisions, while EA has the comparative advantage in rice. Both countries can gain by trading according to their comparative advantages, provided that they exchange the goods in a proportion (called the terms of trade, or international relative price) that lies between their respective relative labour costs. Thus, in our present example, the terms of trade have to be greater than 0.5 and less than 1.2, measured in tons of rice per television.

In the highly simplified world of the comparative advantage model, free trade generally makes all workers better off compared to a situation in which each country tries to make both goods for itself (such self-reliance is referred to as autarky). For example, if the international terms of trade are 3/4 = 0.75 tons of rice per television, then a US worker can obtain more rice by spending 2 hours producing a television and selling it for 0.75 tons of rice, than by growing rice directly (given that a US worker would need 4 hours of labour to produce 1 ton of rice, they could only produce 0.5 tons of rice in 2 hours). Similarly, an EA worker can obtain more televisions by producing and exporting rice than by attempting to produce televisions at home. The increased quantities of the goods that consumers can afford when the goods are produced in the countries where they are relatively cheaper constitute what are called the gains from trade. Of course, it may seem unrealistic in today's world that EA exports rice and imports televisions (and the USA

does the opposite), and indeed it is; in later sections of this chapter we will explain why EA will want to develop its television industry and reverse this pattern of trade (that is, export televisions and import rice) and how it can achieve this.

Changes in the terms of trade

Even in the orthodox approach to international trade, there are some important caveats regarding the gains that countries receive if trade follows comparative advantage. One important qualification is that the gains from trade can be redistributed between countries if the terms of trade shift in favour of one country's exports and against the other's. In our example, if the terms of trade increase from 0.75 to 1 ton of rice per television, then – although both countries are still better off with trade compared with autarky (no trade) – the television-exporting country (USA) gains relatively more and the television-importing country (EA) gains relatively less from the trade. Using our previous example, if a US worker who spends 2 hours producing a television can sell it for 1 ton of rice instead of 0.75 tons of rice, then clearly that worker will be able to consume more rice. However, an EA worker who spends 5 hours growing a ton of rice will be able to purchase less televisions if they only get 1 television per ton of rice instead of 1.33 = 1/0.75 (note that the terms of trade for rice are always the reciprocal of the terms of trade for televisions).

Going beyond this simple example, changes in the terms of trade can have a powerful impact on the national income of countries that export various types of goods in the real world. One important case is the terms of trade for exports of primary commodities, namely, the agricultural and mineral products (including petroleum) that are the main exports of many countries in various regions of the world (for example, Argentina, Mongolia, Nigeria, and Russia). Figure 15.2 shows two key indexes of these international terms of trade, one for crude oil and one for non-fuel primary commodities (agricultural products and non-fuel minerals). Both indexes were constructed by deflating (dividing) the nominal price indexes (which reflect current US dollar price levels) for each type of commodity by a US producer price index, on the assumption that exporters of these goods spend their export revenue largely on imports of industrial products that are priced in US dollars in global markets, so the resulting number measures the commodity exporters' ability to buy industrial goods (relative to 100 in the base year of 1981, when the non-fuel data are first available).

The data in Figure 15.2 illustrate how sharply the purchasing power of the exporters of oil and other commodities over industrial goods has varied

Note: The non-fuel index is not available before 1981.

Sources: Bureau of Labor Statistics (2019), International Monetary Fund (2014, 2019a) and author's calculations.

Figure 15.2 World terms of trade for crude oil and non-fuel primary commodities, 1970–2017

during the past several decades. When the terms of trade for primary commodity exporters rise (as in the 1970s or early 2000s), the real income of the exporting countries (for example, Kazakhstan or Saudi Arabia) rises, but there is a corresponding decrease in the ability of the countries that import these goods (for example, Japan or the European Union) to purchase them and hence a decline in the real income of the latter countries. Evidently, oil prices have fluctuated even more widely than prices of non-fuel commodities, but all commodity exports are subject to destabilizing swings in their terms of trade.

At various points in the twentieth century, many economists in developing countries feared that there would be a long-term decline in the terms of trade for primary commodity exports, which would reduce those countries' ability to import manufactured goods from industrialized countries. As Figure 15.2 shows, although these terms of trade did decline in the 1980s and 1990s, they were generally on a rising trend between 2000 and 2013, owing to a combination of strong global demand (especially from China)

and speculative pressures in global commodity markets. However, the prices of many primary commodities (including oil) collapsed starting around 2014 as a result of weakening demand from producers of industrial goods (for example, the European Union and China). Although a long-term trend is thus difficult to discern, what is clear is that exporters of primary commodities are subject to wild gyrations in their terms of trade that can foster boom-bust cycles (alternating periods of rapid growth and sudden decline or crisis) in their economies.

Although it may seem that it is always better for a country to have high terms of trade for its exports, high terms of trade for primary commodities can be a mixed blessing. In fact, some have claimed that they can become a curse. One reason is that high export prices for primary commodities give countries less incentives to invest in industry and manufacturing, which generally offer greater opportunities for improving technology and raising productivity and living standards in the long run than agriculture or mining. Thus, even if a country enjoys high per capita income as a result of high prices of primary commodity exports for some period of time, the resulting deindustrialization may inhibit the country's long-run development prospects. Other reasons include the destabilizing effects of the inevitable declines in the commodity terms of trade that bring commodity booms to an end, and the environmental degradation that can result from deforestation and overexploitation of soils and mines.

The deindustrializing effects of high terms of trade for primary commodities can be worsened if the increased export revenue leads to the appreciation of the country's currency, which makes its manufacturing industries less competitive in global markets. This phenomenon is often called 'Dutch disease', because it was observed in the Netherlands after the 1959 discovery of natural gas reserves there led to a rise in the exchange rate of the Dutch guilder. Moreover, fluctuations in commodity prices and currency values are often driven by financial speculation, which can be destabilizing for commodity exporters and importers alike. A good example of this problem is found in the Brazilian economy, where the boom in commodity prices in the early 2000s contributed to an appreciation of the Brazilian currency (the real) that in turn led to a fall in the manufacturing share of Brazil's total exports of goods from about 55 per cent in the late 1990s to 35 per cent in 2012 (data from World Bank, 2019). Although the appreciation of the real was later reversed, the damage to the Brazilian manufacturing sector has persisted, with the manufacturing export share barely reaching 38 per cent by 2017.

Distributional consequences of trade (winners and losers)

A second, very important set of qualifications within the orthodox approach concerns the fact that international trade can redistribute income among different groups within a country; that is, some people gain but others may lose as a result of an opening to trade. In reality, the citizens of a country are not all generic 'workers' as in the simple Ricardian model of comparative advantage discussed earlier.[2] There are several dimensions along which the gains and losses from trade may be felt.

At the most basic level, we may distinguish between producers and consumers of any good. In general, consumers benefit if the goods they consume fall in price owing to greater competition from imports (think of cheap imported clothing from South Asia or Central America in the USA or European Union), but domestic producers of these same goods in the importing countries lose (they get lower prices for what they sell, and workers in these industries lose jobs). By the same logic, producers gain but domestic consumers lose when prices of exported goods increase because of high foreign demand for these goods.

However, we can also think of producers as being divided into groups or social classes that own different productive inputs, such as labour, land (natural resources) or capital. In this case, according to the neoclassical Heckscher–Ohlin model of trade, each country will have a comparative advantage in the goods that are relatively intensive in the use of the factors of production (inputs) that are relatively abundant in that country. For example, land-abundant countries such as Australia, Brazil and Canada will export resource-intensive primary commodities, while labour-abundant countries such as Bangladesh and Vietnam will export labour-intensive manufactures (or services, as in the case of India). An extension of this model called the Stolper–Samuelson theorem tells us that owners of the inputs (factors of production) that are relatively abundant in the country gain from free trade, while owners of relatively scarce inputs lose, because the demand for the abundant factor rises when production of the exported good increases, while the demand for the scarce factor falls as the country shifts resources out of the industry producing the imported good and into the export industry.[3] For example, low-skilled US workers are hurt by imports of labour-intensive goods such as apparel or toys, which reduce demand for low-skilled labour in the USA, while high-skilled workers (scientists and engineers) benefit from exports of high-technology goods such as jet aeroplanes and computer software that increase demand for the services of such workers.

The Stolper–Samuelson theorem thus leads us to expect that trade would foster greater inequality between high- and low-skilled workers in advanced countries such as the USA, but would reduce this inequality in developing countries such as Mexico that export goods which are intensive in low-skilled labour. In reality, however, many studies have found that inequality between different groups of workers has increased as a result of trade not only in advanced countries such as the USA, but also in developing countries such as Mexico. This finding has led to the development of new theories that explain how trade can foster greater inequality in all countries, not just in the more advanced ones where low-skilled labour is scarce (Box 15.1). Nevertheless, workers are not always losers from trade. In the countries that have been most successful at export-led growth, such as South Korea and China, real

BOX 15.1

OFFSHORING, TRADE AND WAGE INEQUALITY

Feenstra and Hanson (1997) analysed how the offshoring of jobs from a richer, high-wage country to a poorer, low-wage country can increase inequality among workers in both countries. According to Feenstra and Hanson, the jobs that are offshored (for example, jobs in automobile plants that move from the USA to Mexico) are relatively low-skilled in the richer country but relatively high-skilled in the poorer one. As a result, the average composition of employment shifts toward relatively higher-skilled labour in both countries, so labour demand shifts in favour of the high-skilled workers and wage inequality (the gap between the better-paid, high-skill workers and the lower-paid, low-skill workers) worsens in both countries.

The new models of trade with 'heterogeneous firms', that is, where some firms are more efficient or productive (and hence have lower costs) than others, originally developed by Melitz (2003), can also help to explain why trade tends to worsen inequality. When countries lower their trade barriers, only the most productive (lowest cost) firms will be successful in exporting (or competing with imports) while less productive (higher cost) firms will be unable to compete in global markets (or domestically) and will exit (go out of business). As a result, the average productivity of the remaining firms will increase while industrial employment may shrink (because less workers are needed in the now more productive industries) after a country liberalizes its trade, as found by Trefler (2004) for the case of Canada. In addition, thanks to the lower costs combined with significant market power of the remaining firms, their average profit margins are likely to increase, leading to a higher share of revenue (value-added) going to capital (profits) instead of labour (wages and salaries) (see De Loecker and Warzynski, 2012). Also, because the more productive firms that succeed in exporting tend to use relatively more high-skilled labour, high-skilled workers are likely to benefit more than low-skilled workers (who may actually lose) when trade opens up (see Amiti and Davis, 2012; Egger and Kreickemeier, 2012).

wages of industrial workers have risen significantly in recent decades, causing some of the most labour-intensive industries such as apparel to leave those countries and relocate to lower-wage areas (for example, South Asia and Central America).

The heterodox alternative: imbalanced trade, unemployment and absolute competitive advantages[4]

The traditional theory of comparative advantage rests on the twin assumptions of balanced trade (the value of exports equals the value of imports) and full employment (which implies that everyone who loses a job as a result of imports gets a job in export production or another domestic industry). These assumptions are essential for the validity of the theory: together they ensure that trade promotes efficiency in the allocation of resources and has no effect on the level of employment of the resources. In reality, however, most countries have imbalanced trade: either a surplus, meaning that the value of exports exceeds the value of imports; or a deficit, which indicates the opposite. In addition, there is typically some unemployment in most economies most of the time, and even if there is temporarily something close to full employment, this cannot be guaranteed to continue. Therefore, the heterodox (post-Keynesian) approach examines the impact of trade in the real world where both imbalanced trade and unemployment are usually found.

If full employment is not guaranteed, then workers who lose jobs because of imports may not find jobs in the export sector or any other domestic activity, as is assumed (quite unrealistically) in the theory of comparative advantage. In the absence of full employment, workers who lose jobs because of imports and do not find employment elsewhere (or who get other jobs only at lower wages) will not share in the consumer gains from cheaper imports.[5] For example, recent studies have verified that increased imports from China into the US market caused huge and persistent job losses and wage reductions that were concentrated in the local areas of the USA where the industries that compete with those imports are located (Autor et al., 2016). By the same token, if a country that has unemployed workers is successful in boosting its exports, it can create more jobs and need not sacrifice output of any other goods in order to produce more exports, so the issue of opportunity cost becomes moot. Such a country can possibly reduce unemployment by protecting domestic industries from imports, for example through a tariff (tax on imports) or quota (quantitative restriction on imports), but this will only succeed in raising employment if other countries do not retaliate by levying tariffs or restrictions of their own on the country's exports.

As explained in Chapter 4, modern economies are best described as 'monetary economies of production', which means fundamentally that goods are exchanged for money and not bartered for other goods. The model of comparative advantage, in contrast, rests upon a barter vision of trade: one good (televisions) is exchanged directly for another good (rice), without any role for money. In the real world, however, goods are usually exchanged internationally for major or hard currencies (for example, dollars, euros or yen), which in turn can be converted into other financial assets (for example, bank deposits, Treasury bills or long-term bonds). As a result, some countries can build up trade surpluses by selling more in exports than they buy in imports, thereby acquiring net financial assets from other countries in exchange for their excess exports; while other countries have trade deficits, in which case they must be selling financial assets (or acquiring international debt) to cover the excess of their imports over their exports. For example, when China sells more goods to the USA than it imports from the USA, China uses the excess dollar earnings to increase its holdings of US financial assets such as stocks, bonds and Treasury bills; this in turn increases the US net international debt to China.

Global trade imbalances

Figure 15.3 shows the countries that had the largest trade imbalances (surpluses or deficits in excess of US$25 billion, measured by current account balances) as of 2018. The USA had by far the world's largest deficit, followed by a few other Anglophone countries (the United Kingdom, Canada and Australia) and several emerging market nations (India, Indonesia, Turkey and Argentina). The largest surpluses are found in some manufacturing exporters, such as Germany, Japan and Taiwan, and major resource exporters, such as Russia and Saudi Arabia (some surplus countries, such as the Netherlands, fit both descriptions). Of course, many other countries have deficits or surpluses that are large percentages of their own GDP, but the ones shown in this figure loom largest in terms of global financial flows (since surplus countries must be net lenders, while deficit countries are net borrowers).

One of the factors that affects trade balances, especially for countries that export manufactured goods, is a country's unit labour costs (wage costs per unit of output) compared to other countries' unit labour costs, measured at the prevailing exchange rate (price of one currency in terms of another).[6] A country with lower unit labour costs, converted to a foreign currency (such as the US dollar), will be likely to export more and import less, thereby tending to give it a bigger trade surplus. For example, countries such as China and Mexico have taken advantage of relatively low wages combined with

Source: International Monetary Fund (2019b). Data for several countries are International Monetary Fund (IMF) staff estimates.

Figure 15.3 Countries with current account surpluses and deficits in excess of US$25 billion, 2018

rising productivity in their export industries to achieve competitive advantages in many manufactured products, such as electronics and automobiles. In addition, in the late 1990s and early 2000s, China also maintained an artificially low value of its currency, the renminbi (also called the yuan), thereby making its exports even cheaper in dollar terms than they would have been otherwise. Indeed, as the renminbi has appreciated (risen in value) and Chinese wages have increased since around 2005, China's trade surplus has decreased and some labour-intensive industries (for example, apparel) have left China for countries that have even lower wages (for example, Cambodia or Vietnam). By 2018, the year shown in Figure 15.3, China no longer had one of the world's largest current account surpluses, as it had in earlier years.[7]

In the euro area, where there is no internal exchange rate (because all countries use the same currency), one country can still increase its competitive advantages over other users of the same currency by making its nominal wages (in euros per hour) grow more slowly relative to the productivity of its

workers; as, for example, Germany did relative to Italy and Spain in the decade prior to 2008.[8] By the same token, euro area countries that are struggling to emerge from an economic crisis may be forced to suppress their wages relative to their productivity – a process sometimes called internal devaluation – as, for example, Greece, Ireland and Spain have done since 2009.

Whenever some countries have large trade surpluses and others have large deficits, much of global trade is following absolute competitive advantages (that is, the lowest monetary cost of production at prevailing exchange rates) rather than comparative advantages.[9] In this situation, international trade is not generally fulfilling its supposed mission of making sure that goods are always exported by the countries that can produce them most efficiently; rather, some countries are exporting goods that they would not export if trade were balanced, while other countries are importing those goods instead of producing them for themselves.

Trade and unemployment

The issue of imbalanced trade is also linked to the problem of unemployment. For example, in the case of the world's largest bilateral trade imbalance – the US deficit with China – estimates show that US job losses (either total, or just in manufacturing) attributable to trade with China since the early 2000s could be of the order of 2.4 million or higher.[10] Deficit countries, however, do not always have high rates of unemployment. Sometimes, a country can have both a large trade deficit and a reduced unemployment rate, but this usually occurs when the country is borrowing to sustain high aggregate demand at home. In this case, a country's consumers can buy goods and services in excess of what their current income would otherwise permit, thereby boosting the country's demand for imports and lowering its trade balance (as in the USA or Spain in the early 2000s, when both countries were experiencing debt-driven housing booms). In such situations, the borrowing country can temporarily sustain low overall unemployment along with a trade deficit, often by creating jobs in low-wage, non-tradable sectors (such as services and construction) instead of high-wage, tradable goods industries (principally manufacturing); which is exactly how the USA managed to prevent overall unemployment from rising in spite of the job losses associated with Chinese imports. Further, such a debt-led boom usually ends in a financial crisis and recession – during and after which unemployment soars – as occurred in the USA in 2007–08 and in Greece, Portugal, Spain and other 'peripheral' euro area countries in 2009–12.

After a financial crisis, deficit countries that had enjoyed temporary booms based on borrowing are usually compelled to make painful adjustments

through higher unemployment and reduced incomes that in turn depress their purchases of imports, thereby improving their trade balances. For example, Greece and Spain do not appear in Figure 15.3 because by 2018 they had eliminated their previously large trade deficits from before the euro area crisis of 2009–12, but during the crisis years and the early stages of their recoveries their unemployment rates soared to around 20–25 per cent as a result of fiscal austerity and deep recessions.

Long-run development and infant-industry protection

Even if we abstract from the difficulties of unemployment and imbalanced trade, the traditional theory of comparative advantage suffers from another defect. The traditional theory is static, which means that it investigates only what is the best option for a country at a given moment in time, based on its current resources and technology. However, in the long run, countries must improve their technology in order to grow and raise their living standards, and in the process they may need to develop the capability to produce goods in which they do not have a pre-existing comparative advantage.

In the example shown in Table 15.1, adopting free trade and specializing in rice production for export was the best option for EA only because EA's technology for producing televisions was so poor. If EA could improve its technology, it could potentially become a manufacturing region that would export goods such as televisions to the USA, instead of the other way around. To see this possibility, consider the example shown in Table 15.2, where EA has lowered its labour cost to 2 hours per television, the same as in the USA; perhaps, at least initially, by importing or imitating US television technology. Now, the relative labour cost of a television is lower in EA than in the USA (2/5 = 0.4 versus 2/4 = 0.5), so EA switches to having a comparative advantage in televisions. Henceforth, EA will export televisions in exchange for rice imported from the USA. This can be called 'dynamic comparative advantage', because over time the less developed country acquires the technical know-how required to produce a more advanced product and changes the direction of its trade.

Table 15.2 Dynamic comparative advantage: technological improvement in East Asia

	Hours of labour required to produce:	
	Televisions (per unit)	Rice (per ton)
United States	2	4
East Asia	2	5

But, how could the EA television industry ever get off the ground? If EA allows free trade under the conditions shown in Table 15.1, it would remain permanently specialized in rice and its manufacturing industries would remain underdeveloped; any potential television producers could not compete with relatively cheaper US imports. Therefore, EA may need to use tariffs (taxes on imports) or quotas (quantitative restrictions) to protect local producers of televisions from imports until those producers become efficient enough to compete on their own, a policy that is known as infant-industry protection. Such protection is designed to enable a domestic industry to acquire improved technology and lower its costs to the point at which, eventually, the industry can survive without protection (and may possibly start to export).

There are some important qualifications to the case for infant-industry protection. First, in addition to (or instead of) trade policies, a government may choose to utilize other types of industrial policies. For example, the government could subsidize an industry through various means, such as by offering low interest rates on loans or providing necessary infrastructure and training. Thus, the promotion of infant industries need not be done exclusively through trade protection. Second, if trade protection is used, the tariffs and quotas should not be too restrictive or permanent; otherwise, the protected industry might not have incentives to become more efficient and might never be able to export. There are cases of countries that have protected and subsidized infant industries that, in effect, never grew up to become efficient and competitive exporters (for example, Brazilian computers). This is why many developing countries, such as Mexico and India, have reduced formerly high levels of protection and sought to expose their producers to greater global competition. Protecting domestic industries via tariffs or quotas does raise the costs of their products for domestic consumers, so it is important that the long-term gains from increased efficiency and eventual exports should outweigh the short-term costs of the protection.

Third, governments seeking to promote new industries face a choice of allowing foreign multinational corporations (MNCs) to produce the more advanced goods, or else promoting domestic companies (sometimes called 'national champions') that could potentially produce them instead. The former path is in many ways easier, because foreign MNCs already have the necessary technology and know-how, but the MNCs may not want to share the technology and know-how with the host country. Therefore, even though it is more difficult, it may be more beneficial in the long run to try to promote national firms in order to enhance domestic technological learning. Alternatively, it may be possible to negotiate with foreign MNCs to ensure that they share their technology, for example by encouraging them to train

domestic workers in advanced technology or to form partnerships with local firms (as China has done). Most major countries that have successfully developed in the past (including the USA, Germany and Japan) have used infant-industry protection and other industrial policies at critical points in their history, and the few countries that have made the leap from developing to industrialized in recent decades (for example, South Korea) have done so by enabling their own national companies (champions) to become global innovators and technological leaders (the Korean multinational Samsung is a good example) (see Chang, 2002; Lee, 2013).

Trade liberalization, trade agreements and trade wars

Why countries join trade agreements

As noted above, it can sometimes be beneficial for countries to use restrictive trade policies strategically to help develop their industries. In addition, many countries adopt protectionist policies with the intention of boosting domestic employment, or simply to protect domestic producers (firms, workers and/or farmers) who could lose sales or jobs if imports were liberalized. Nevertheless, all of these policies – regardless of the motivation – are likely to have adverse consequences for other nations. The protection of an industry in one country can mean a loss of exports and jobs for another country. Similarly, a subsidy used to promote exports in one country can result in artificially cheap imports that displace domestic producers and reduce employment in another. To use our previous example, if EA protects or subsidizes its television industry, this will cause job losses for US television workers, who are unlikely to become rice farmers and at best are likely to be re-employed at lower wages in the service sector.

For these reasons, a long line of economists going back to Adam Smith (1776/1976) has referred to protectionism as a 'beggar-my-neighbour' policy: even if it enriches the protecting nation, it may impoverish others. Moreover, the fact that one country's trade and industrial policies can adversely affect other countries' industries and employment may lead to retaliation: a country may impose tariffs or other protectionist measures in response to a foreign country's similar interventions. If a large number of countries adopt retaliatory trade barriers at the same time, the effect is to shrink global trade and lessen the gains that any country can derive from participation in the world market. Whereas Smith and later orthodox economists emphasized that protectionism and retaliation would lessen efficiency in resource allocation, leading to higher costs for consumers, post-Keynesian economists led by Joan Robinson (1947) emphasized that such policies – if

adopted simultaneously by too many major countries – would reduce global aggregate demand and worsen unemployment.

Historically, this problem became most acute during the Great Depression of the 1930s. Although high tariffs and frequent retaliation did not cause the depression, as some have claimed, they certainly did not help the global economy to recover. Therefore, after World War II the victorious allies began a process of multilateral trade liberalization by engaging in broad-ranging negotiations over reciprocal reductions in tariffs and other trade barriers. This process started with the General Agreement on Tariffs and Trade (GATT), originally launched at a Geneva conference in 1947, and continued through the formation of the World Trade Organization (WTO) in 1995. As of September 2019, the WTO had 164 member countries.[11]

In addition, smaller groups of countries have formed what are variously known as bilateral, regional or preferential trade agreements. Some prominent examples include the European Union (EU) and the United States–Mexico–Canada Agreement (USMCA, formerly the North American Free Trade Agreement or NAFTA). The simplest type of trade agreement is a free trade area, in which trade barriers among the member countries are reduced or eliminated. More ambitious efforts at economic integration include a customs union (in which the member countries also adopt common external tariffs), a common (or single) market (in which the members allow free flows of labour and capital as well as goods and services), and an economic union (in which, in addition to all of the above, the countries seek to harmonize other aspects of their economic policies). Another type of economic integration scheme, a monetary union, is discussed in Chapter 17 for the case of the euro area.

However, one should be cautious in assessing actual trade agreements, because they do not necessarily live up to their names. For example, Mercosur is supposed to be the 'Common Market of the South', but the member countries (Argentina, Brazil, Paraguay and Uruguay) have never fully eliminated tariffs and other trade barriers on each other's goods or negotiated a common external tariff, so it is not even really a free trade area or customs union. USMCA does not guarantee pure free trade; for example, it allows the member countries to impose tariffs on each other in certain circumstances.

The orthodox analysis of trade agreements emphasizes the issue of trade creation versus trade diversion: do the efficiency gains (cost reductions) from increasing trade among the member countries outweigh the possible losses from trading less with outside nations, which might be able to supply imports

at still lower costs? However, trade agreements have often been formed for many reasons that go beyond standard calculations of the net gains from trade creation and diversion. Other economic reasons for trade agreements include: wanting preferential access for goods exported to a major trading partner's market; taking advantage of scale economies by producing for a larger market area; trying to attract MNCs to invest in a country because it is a member of a larger trading bloc; and (in the other direction) opening up foreign countries to investments by the home country's MNCs. The first three of these reasons have been important motivations for smaller or less developed countries that have joined integration schemes with larger or more advanced countries, such as Mexico in NAFTA/USMCA and Ireland in the EU. Sometimes the motivation is political, as when the countries in a region or alliance want to foster greater integration of their societies and be more unified in response to perceived external threats; this was an important reason for the efforts at European integration that led to the formation of the EU in 1992 and its subsequent expansion (see also Chapter 17).

Regional or preferential trade agreements gained a new impetus as a result of the breakdown of multilateral trade negotiations after 2001. A new WTO round was launched at Doha in 2001, but negotiations subsequently ceased and at the time of writing (late 2020) there was no prospect of their resumption. The major players such as the USA, EU and leading emerging market nations (Brazil, China, India, and so on) could not agree on various outstanding issues, including US demands for strengthened 'intellectual property rights' (more accurately, extended patent and copyright protection that increases US firms' monopoly power) and the unwillingness of many countries (from the EU to Japan) to further liberalize their agricultural sectors. As the multilateral process became stuck, many countries went ahead and formed more limited trade agreements with varying numbers of partners. In particular, the USA used its leverage over smaller countries (for example, in Central America and the Dominican Republic) that wanted access to its market to win provisions that it could not get in the WTO process, such as strengthened intellectual property rights (patents, copyrights, and so on) and investor–state dispute settlement (ISDS) procedures that favour multinational corporations.

As Rodrik (2018, p. 75) has written:

> Contemporary trade agreements go much beyond traditional trade restrictions at the border. They cover regulatory standards, health and safety rules, investment, banking and finance, intellectual property, labor, the environment, and many other subjects. They reach well beyond national borders and seek deep integration among nations rather than shallow integration.

In other words, what modern trade agreements do is not only to write the rules governing trade itself, but also to limit the options for national governments in other policy dimensions (for example, by prohibiting controls on international capital flows or inhibiting environmental regulations that would impinge on corporate profits). Typically, trade negotiations are heavily influenced by lobbyists representing corporate and financial interests, much more than any other social interests, and there is limited democratic input into what are usually very secretive deliberations. As a result, it should not be surprising that the trade negotiating agenda has been largely captured by those wealthy and powerful interests:

> [contemporary] trade agreements are shaped largely by rent-seeking [monopoly-creating], self-interested behavior on the export side. Rather than reining in protectionists, trade agreements empower another set of special interests and politically well-connected firms, such as international banks, pharmaceutical companies, and multinational corporations. Such agreements may result in freer, mutually beneficial trade, through exchange of market access. But they are as likely to produce welfare-reducing, or purely redistributive outcomes under the guise of free trade. (Rodrik, 2018, pp. 275–6)

The new economic nationalism

By the second decade of the twenty-first century, a backlash against such trade agreements and economic integration more broadly (including labour migration) began to take hold in several leading countries. Two dramatic events that marked the rise of this new economic nationalism were the victory of the 'leave' vote in the Brexit referendum in the United Kingdom (UK) and the election of Donald J. Trump as US President (the latter without a majority in the popular vote, but a win in the US Electoral College system), both in 2016. 'Brexit' refers to British exit from the EU, while Trump vowed to 'make America great again' by enacting protectionist policies (and blocking immigration).

This new economic nationalism had its roots in popular responses to the job losses and heightened inequality that, as discussed earlier, occurred partly as a result of increased openness to international trade and investment. Although other forces (such as technological changes and domestic deregulation) also contributed to these disruptions, popular attention often focused on trade agreements (such as NAFTA and the EU) or trade imbalances (such as the large bilateral US deficit with China). So-called populist political leaders (critics might call them demagogues) stoked these sentiments and transformed trading partners or EU bureaucrats – along with immigrants

and minorities – into scapegoats for their countries' economic problems, real or perceived, ignoring more close-to-home causes such as fiscal austerity policies, increased monopoly power and inadequate public investment. These leaders promised that withdrawing from, or renegotiating, various trade agreements would help to spark recoveries of industries, jobs and wages in their countries. Critics countered that these policies were more likely to cause additional dislocations and hardships, without bringing the promised gains or reversing the losses of the past.

At the end of 2020, British Prime Minister Boris Johnson negotiated a post-Brexit trade agreement with the EU in which there would be no tariffs on UK-EU trade in goods, but services were not covered and administrative costs would rise as a result of new customs requirements (except between Northern Ireland and Ireland). Since the UK has left the EU customs union permanently, Brexit allows the UK to negotiate new trade agreements with other countries such as the USA, which the UK could not do as long as it remained within the EU (which negotiates trade agreements on behalf of all member countries). However, it is doubtful that an economically independent UK – as a medium-sized economy – will be in a strong position to win advantageous concessions from the USA, EU, China, or other major trading partners. More likely, these larger and more powerful entities will instead exact more concessions from the UK in any future agreements. The UK government also seeks greater regulatory autonomy and the ability to restrict immigration from EU countries through Brexit; however, the UK-EU trade agreement limits how much British regulations can differ from EU regulations.

On the other side of the Atlantic, one of Trump's first actions after he took office in January 2017 was to withdraw the USA from the Trans-Pacific Partnership (TPP), a trade and investment agreement with 11 other Pacific Rim countries that had been negotiated by his predecessor (Barack Obama) but was never approved by the US Congress. Subsequently, the other 11 countries (including Japan, Canada, Mexico and Australia) went ahead and approved a modified version of the agreement, called the Comprehensive and Progressive Agreement for TPP (CPTPP or TPP-11), in which they omitted some of the more controversial US-backed provisions (such as extended patents for pharmaceuticals and ISDS protections for foreign investors). Then, Trump launched what came to be called his 'trade war', by levying tariffs on various imported goods and starting a major battle with China over its trade and investment policies, as well as by renegotiating NAFTA (Box 15.2).

BOX 15.2

DONALD TRUMP'S TRADE WARS

After withdrawing the USA from the TPP (as discussed in the text) early in 2017, former President Trump then launched a set of more aggressive trade moves, including the imposition of tariffs on imports of steel and aluminium products (on 'national security' grounds) and hundreds of billions of US dollars worth of imports from China (based on a US law that grants the President broad discretion to take actions against countries that allegedly engage in unfair trade practices). One small country (South Korea) quickly negotiated a managed trade agreement, which limited its steel exports to the USA, to evade Trump's wrath. But the larger and more powerful countries, such as China and the EU, instead retaliated with tariffs of their own on US exports (for example, Harley-Davidson motorcycles). Meanwhile, the tariffs backfired in several respects. The steel and aluminium tariffs hurt downstream US producers who use these goods as inputs, while the tariffs on Chinese imports mostly induced firms to source the imported goods from other low-wage countries (such as Mexico or Vietnam) instead of producing them in the USA. US farmers suffered significant losses in revenue and prices as their export markets for products such as soybeans collapsed as a result of China's retaliation.

Ostensibly, the tariffs on China were supposed to induce that country to negotiate in response to a series of US demands (for example, ending its alleged 'theft' of foreign firms' intellectual property), but China was reluctant to make fundamental changes to its (highly successful) economic policies and development model – or even to make more modest concessions – in response to US threats. Furthermore, the uncertainty caused by Trump's capricious way of announcing (and sometimes withdrawing) tariffs based on his personal whims and political fortunes inhibited business investment in the USA and elsewhere. In this atmosphere, nothing remotely resembling a grand renaissance of US manufacturing industries could be observed: not in steel and aluminium, nor in any other industries.

In what was perhaps its most successful move, the Trump administration launched a renegotiation of NAFTA with Canada and Mexico (by threatening to withdraw the USA from the agreement if it was not rewritten). The new accord was labelled the 'US–Mexico–Canada Agreement' (USMCA), deliberately eschewing any mention of 'North America' or 'free trade'. USMCA actually leaves much of the structure of NAFTA intact, but makes changes in certain key areas. It contains greater safeguards for labour rights – especially the right to organize free and independent unions – compared to NAFTA, and more in line with what had been included in TPP. These provisions were strengthened in response to demands from Democrats in the US Congress as a condition for their support. USMCA weakens or abolishes the ISDS regime for Canada and most sectors of the Mexican economy (which Trump's US Trade Representative Robert Lighthizer believed would lessen incentives for foreign investment in Mexico).

USMCA increases the regional North American content requirement ('rule of

➡

origin') for automobiles from 62.5 to 75 per cent, and requires that about 40 per cent of the value of a car must be produced by labour earning at least US$16 per hour, in order for a car to be sold duty-free. Lighthizer hoped this provision would induce the relocation of many automotive jobs back to the USA, but instead some auto producers have announced plans to raise wages in their Mexican plants. USMCA enacts rules about e-commerce that favour large US technology companies. USMCA also contains a 'sunset clause', which requires its further renegotiation starting six years after its enactment, or the agreement will be suspended after 16 years (another effort by Lighthizer to increase uncertainty for foreign investors in Mexico). Clearly, it is a mixed bag, and overall USMCA is unlikely to create significant economic benefits. Nevertheless, the agreement was approved by all three member countries and went into effect on 1 July 2020.

The new administration of President Joe Biden will have to decide how to unwind Trump's trade wars while still pursuing what it perceives to be US interests in global trade negotiations. Early indications are that Biden will not remove the tariffs on China quickly, and will seek to negotiate over those only after repairing relations with US allies and focusing on domestic job creation first. The desire to assuage US allies will probably lead to the elimination of the steel and aluminium tariffs and renewed US participation in the WTO dispute settlement mechanism. USMCA seems like settled policy for the present, but the Biden administration will put pressure on Mexico to uphold the labour rights provisions in the agreement. It is not clear if the Biden administration will want to rejoin the (renamed) CPTPP, but one way or another the administration is likely to focus on new strategies for confronting the rise of Chinese influence in the Asia-Pacific region, including the emergence of a Chinese-led trading bloc in the newly formed Regional Comprehensive Economic Partnership (RCEP).

Only time will tell whether this new economic nationalism – combined with the after-effects of the covid-19 pandemic – will continue to disrupt the global economy. However, it is important not to think that the world should just go back to the *status quo ante*, when trade agreements combined with other neoliberal policies created their own kinds of economic dislocations and inequalities. As de Grauwe (2016) has argued, it is perhaps time to put a moratorium on new trade agreements 'as long as we do not keep in check the environmental costs generated by free trade agreements and as long as we do not compensate the losers of globalisation'. And perhaps it is time to return to multilateral trade negotiations, but with a new emphasis on allowing different types of policy regimes and development strategies to flourish, while safeguarding the rights of workers and consumers, protecting the environment from further devastation, and using coordinated Keynesian macroeconomic policies to ensure that full employment is maintained globally. Last but not least, reform of the international monetary system (including

the removal of the US dollar from its role as the chief reserve currency) could help to reduce trade imbalances and defuse trade tensions, while making sure that any new system would require the surplus countries to bear a significant part of the adjustment burden by adopting expansionary macroeconomic policies, as advocated long ago by Keynes (a proposal revived by Davidson, 1992–93).

Manufactured exports, the fallacy of composition and global value chains

In spite of the controversy over trade liberalization and trade agreements, it is true that countries that have been successful at exporting manufactured goods have generally (with some exceptions, discussed below) had rapid economic growth that has significantly raised their income levels. Why are exports of manufactured goods usually so beneficial for promoting economic development? One key reason is economies of scale: by producing for a larger market, exporting firms can reduce their average costs by spreading out the fixed costs (for example, machinery and equipment, or research and development) over a greater quantity of output; this is more likely to occur in manufacturing than agriculture or services. A second reason is that manufacturing activities engender greater technological learning within firms and more spillovers of knowledge between different firms and industries. Third, manufacturing generally offers higher productivity of labour (value-added per worker) than most other sectors, which implies that transferring labour from traditional agricultural or informal services into manufacturing is likely to increase average productivity in the economy as a whole. Fourth, manufacturing has advantageous demand-side characteristics: many manufactured goods have high income elasticities, which implies that the demand for them grows rapidly as incomes rise in global export markets. (For further discussion of the importance of income elasticities, see Chapter 16.)

In spite of these attractions of a development strategy focused on manufactured exports, only a relatively small number of developing countries (mostly in East Asia) have succeeded in using manufactured exports to leverage rapid, sustained growth since the latter half of the twentieth century. For countries that are abundant in natural resources (especially in the Middle East, South America and Africa), the strong pull of comparative advantage has led them to continue to specialize in primary commodity exports (in some cases, reverting after previously attempting to industrialize through infant-industry protection). Also, even countries that could be more successful in manufactures sometimes fail to adopt policies that would be conducive to promoting manufactured exports, for example by allowing their currencies to become

overvalued, not promoting or protecting infant industries, or not investing in the education and infrastructure required for more advanced industries to thrive.

But among those countries that have attempted to pursue manufacturing export-led growth, the number of success stories has been limited by the 'fallacy of composition', or adding-up constraint (Blecker and Razmi, 2010). To wit, not all countries can simultaneously increase their exports of manufactured goods at the very high rates seen in the most successful cases (for example, Korea and Taiwan in the 1970s–80s, or China in the early 2000s). The markets in the rich, industrialized countries are only so big, and grow more slowly than world trade (recall that global trade grew faster than global GDP up to 2008, as shown in Figure 15.1). Once the domestic producers in a given sector in an industrialized country (for example, local manufacturers of textiles and apparel) have been displaced, then the developing countries that want to enter these markets must compete very intensively for export opportunities in those markets.

Only the countries with the lowest monetary costs of production, the most competitive exchange rates and other favourable conditions can succeed, while the others are likely to fall behind and fail to achieve the desired export-led growth. Some countries may escape from this trap by moving 'up the industrial ladder' to more high-technology products, and may even become technological innovators themselves; Korea and China are good examples. In the end, however, not all countries can succeed in enjoying rapid export-led growth at the same time, especially if they seek to specialize in similar types of export products. The very success of some countries thus creates challenges, if not defeats, for others. For example, Mexico displaced Japan, Korea and Taiwan in many segments of the US market after liberalizing its trade and joining NAFTA in the late 1980s and 1990s, but subsequently China displaced Mexico after it joined the WTO in 2001 (Blecker and Esquivel, 2013).

The Mexican case suggests another cautionary note: although the total value of its exports of manufactures has grown enormously, a large portion of these exports consists of goods that are merely assembled in the country, using imported parts and components. As a result, the value-added (the difference between total value and materials costs) is only a small portion of the total (gross) value of many Mexican exports, and therefore the gains in domestic job creation have been disappointing. Furthermore, Mexico's manufactured exports are largely produced by (or under contract with) foreign MNCs that generally keep their most innovative and high value-added operations elsewhere. These elements of Mexico's export strategy, and the

heavy dependency on imported intermediate goods that they have created, have limited the overall growth gains that the country has received from its exports (see Ibarra and Blecker, 2016).

The Mexican example also reveals the importance of a relatively new challenge for developing countries in today's global economy. This is the emergence of global value chains (sometimes also called supply chains), whereby production processes are spread out over many different countries, with different stages of production taking place in (and inputs supplied from) numerous different countries. Thus, an increasing portion of global trade today consists in intermediate and semi-finished products, as opposed to the more traditional raw materials and finished goods. This new pattern of trade is also sometimes referred to as vertical specialization, because countries specialize in different stages of production (for example, labour-intensive assembly operations) rather than complete final goods. As a result of this new pattern of trade, obtaining certain manufacturing operations (for example, production of simple parts and assembly of components) may not create the kinds of 'backward and forward linkages' to upstream and downstream domestic industries that they did in the past. Countries do need to participate in global value chains in order to be players in today's international economy, but emerging nations also need to be careful that they eventually move up into the more innovative and higher value-added links in these chains rather than remaining permanently limited to labour-intensive assembly operations.

Conclusion

Finding the most beneficial trade strategy for a country today is much more complicated than the old (and increasingly sterile) debate about free trade versus protectionism. The countries that have grown most successfully in the long run are those that have strategically deployed prudent means of promoting nascent domestic industries and encouraging exports, especially of manufactures. But each country needs to find its own best strategy, given its economic structure and external constraints.

Paradoxically, the policies that may be most important for making the global trading system work more to the benefit of all nations may not be trade policies or trade agreements at all. Especially, reforming the international monetary system to prevent some countries from achieving trade surpluses based on undervalued currencies could help to restore more balanced and mutually beneficial trade. In addition, competition over global trade and investment opportunities is exacerbated by the chronic lack of aggregate demand throughout much of the world economy, a problem that has been

aggravated by financial crises and fiscal austerity policies. By adopting the types of macroeconomic policies recommended in other parts of this book – especially by targeting monetary and fiscal policies on the achievement of full employment – the global trading system could operate more to the benefit of all countries and with less conflict than it does at present.

NOTES

1 The rising share of trade in global income may be exaggerated to some extent by rising prices of primary commodities, such as oil, in certain periods, as well as the double-counting of imported intermediate goods that are used in export production. Nevertheless, the fact that trade has grown substantially faster than GDP is not disputed.

2 Ricardo's original model of international trade had three classes of economic agents: workers, capitalists and landowners; what has been presented here is the so-called Ricardian trade model as portrayed in standard textbooks. Recall the discussion of Ricardo in Chapter 3 and see Maneschi (1992) and Bernhofen and Brown (2018) for broader perspectives on Ricardo's views about trade.

3 The Heckscher–Ohlin model is named after early twentieth-century Swedish economists Eli Heckscher and Bertil Ohlin; the Stolper–Samuelson theorem is named after US economists Wolfgang Stolper and Paul Samuelson, who published it in 1941. Both the Heckscher–Ohlin model and the Stolper–Samuelson theorem assume that the factors of production are 'mobile' between industries, in the sense that any given unit of labour, land or capital could possibly be employed in producing any of the traded goods.

4 See also Chapter 13 on the role of exports in long-run economic growth, and Chapter 16 on the theory of balance-of-payments constrained growth.

5 Workers who do not find jobs may remain unemployed in advanced countries such as the USA or the UK, where there are social insurance mechanisms to sustain them at least temporarily. In less developed countries, however, such welfare mechanisms often do not exist, and workers who do not find jobs in the modern sector cannot survive long without some sort of work. Thus, such workers tend to end up in the 'informal sector', for example by becoming street vendors or working in small shops, where they have very low productivity and incomes. In these cases, we say that the workers are underemployed instead of unemployed.

6 Note that, by definition, unit labour cost equals WL/PY, where W is the nominal wage rate (money units per hour), L is hours of labour, P is the monetary price of the goods, and Y is the quantity of output. Since this ratio can also be written as $(W/P)/(Y/L)$, we can see that unit labour cost also equals the ratio of the real wage to the productivity of labour (quantity of output per hour). Thus, unit labour costs can be reduced either by suppressing wages or, alternatively, by increasing productivity. All this is in domestic currency; unit labour costs then get translated into a foreign currency by dividing by the exchange rate measured as the domestic currency price of the foreign currency (for example, pesos per dollar in Mexico).

7 China had a current account surplus in excess of US$100 billion every year from 2005 to 2017; its surplus peaked at US$421 billion in 2008 (data from International Monetary Fund, 2019b).

8 See data from the Organisation for Economic Co-operation and Development (OECD, 2019).

9 Note that absolute competitive advantage, in the sense of lowest monetary cost, is not the same as absolute productivity advantage, in the sense of lowest labour time or highest labour productivity. The heterodox theory emphasizes the former, not the latter.

10 For varying estimates of the job losses caused by US trade with China, see Acemoglu et al. (2016), Autor et al. (2016) and Scott (2017).

11 Verified from the WTO website, accessed 10 August 2019 at https://www.wto.org/.

REFERENCES

Acemoglu, D., D.H. Autor, D. Dorn, G.H. Hanson and B. Price (2016), 'Import competition and the great US employment sag of the 2000s', *Journal of Labor Economics*, **34** (S1), S141–98.

Amiti, M. and D.R. Davis (2012), 'Trade, firms, and wages: theory and evidence', *Review of Economic Studies*, **79** (1), 1–36.

Autor, D.H., D. Dorn and G.H. Hanson (2016), 'The China shock: learning from labor-market adjustment to large changes in trade', *Annual Review of Economics*, **8**, 205–40.

Bernhofen, D.M. and J.C. Brown (2018), 'On the genius behind David Ricardo's 1817 formulation of comparative advantage', *Journal of Economic Perspectives*, **32** (4), 227–40.

Blecker, R.A. and G. Esquivel (2013), 'Trade and the development gap', in A. Selee and P. Smith (eds), *Mexico and the United States: The Politics of Partnership*, Boulder, CO: Lynne Rienner, pp. 83–110.

Blecker, R.A. and A. Razmi (2010), 'Export-led growth, real exchange rates and the fallacy of composition', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 379–96.

Bureau of Labor Statistics (2019), 'Producer price indices', accessed 4 August 2019 at www.bls.gov/data.

Chang, H.J. (2002), *Kicking Away the Ladder: Development Strategy in Historical Perspective*, London: Anthem.

Davidson, P. (1992–93), 'Reforming the world's money', *Journal of Post Keynesian Economics*, **15** (2), 153–79.

de Grauwe, P. (2016), 'How far should we push globalisation?', *Social Europe*, 3 November, accessed 8 September 2019 at https://www.socialeurope.eu/far-push-globalisation.

De Loecker, J. and F. Warzynski (2012), 'Markups and firm-level export status', *American Economic Review*, **102** (6), 2437–71.

Egger, H. and U. Kreickemeier (2012), 'Fairness, trade, and inequality', *Journal of International Economics*, **86** (2), 184–96.

Feenstra, R.C. and G.H. Hanson (1997), 'Foreign direct investment and relative wages: evidence from Mexico's maquiladoras', *Journal of International Economics*, **42** (3–4), 371–94.

Ibarra, C.A. and R.A. Blecker (2016), 'Structural change, the real exchange rate and the balance of payments in Mexico, 1960–2012', *Cambridge Journal of Economics*, **40** (2), 507–39.

International Monetary Fund (IMF) (2014), 'International Financial Statistics', database, accessed 4 June 2014 at www.elibrary-data.imf.org.

International Monetary Fund (IMF) (2019a), 'Primary Commodity Price System', database, accessed 4 August 2019 at https://data.imf.org/commodityprices.

International Monetary Fund (IMF) (2019b), 'World Economic Outlook', database, April 2019 edition, accessed 4 August 2019 at www.imf.org/external/pubs/ft/weo/2019/01/weodata/index.aspx.

Lee, K. (2013), *Schumpeterian Analysis of Economic Catch-Up: Knowledge, Path-Creation and the Middle Income Trap*, New York: Cambridge University Press.

Maneschi, A. (1992), 'Ricardo's international trade theory: beyond the comparative cost example', *Cambridge Journal of Economics*, **16** (4), 421–37.

Melitz, Mark J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica*, **71** (6), 1695–1725.

Organisation for Economic Co-operation and Development (OECD) (2019), 'OECD Data, unit labour costs', accessed 7 September 2019 at https://data.oecd.org/lprdty/unit-labour-costs.htm.

Ricardo, D. (1821/1951), *Principles of Political Economy and Taxation*, 3rd edition, reprinted, Cambridge, UK: Cambridge University Press.

Robinson, J. (1947), *Essays in the Theory of Employment*, 2nd edition, Oxford: Basil Blackwell.

Robinson, J. (1956), *The Accumulation of Capital*, London: Macmillan.

Robinson, J. (1962), *Essays in the Theory of Economic Growth*, London: Macmillan.

Rodrik, D. (2018), 'What do trade agreements really do?', *Journal of Economic Perspectives*, **32** (2), 73–90.

Scott, R.E. (2017), 'Growth in US-China trade deficit between 2001 and 2015 cost 3.4 million jobs', Economic Policy Institute, Report, 31 January, accessed 8 September 2019 at https://www.epi.org/publication/growth-in-u-s-china-trade-deficit-between-2001-and-2015-cost-3-4-million-jobs-heres-how-to-rebalance-trade-and-rebuild-american-manufacturing/.

Smith, A. (1776/1976), *An Inquiry into the Nature and Causes of the Wealth of Nations*, Chicago, IL, USA: University of Chicago Press.

Trefler, D. (2004), 'The long and short of the Canada-US Free Trade Agreement', *American Economic Review*, **94** (4), 870–95.

World Bank (2019), 'World Development Indicators Databank', accessed 4 August 2019 at https://databank.worldbank.org/source/world-development-indicators.

A PORTRAIT OF JOAN ROBINSON (1903–83)

Joan Robinson was a colleague of John Maynard Keynes at the University of Cambridge, UK, and a key member of his intellectual circle in the 1930s, when he was writing *The General Theory*. She helped to bring the ideas of Karl Marx and Michał Kalecki into the somewhat insular world of Cambridge, and was also influenced by her discussions and debates with fellow Cantabrigians (especially Piero Sraffa and Nicholas Kaldor). She made contributions to economics in a wide range of areas including imperfect competition, endogenous money and economic development. However, she is best known for her work on economic growth and income distribution, especially in her *Accumulation of Capital* (1956) and *Essays in the Theory of Economic Growth* (1962). She was a major protagonist in the Cambridge controversies on capital theory from the 1950s to the 1970s, which showed that the conventional idea of an inverse relationship between the 'quantity' of capital and the rate of profit (or rate of interest or returns to capital) is fundamentally flawed.

Robinson was especially important in bringing income distribution into a central place in the theory of economic growth. Her own model of capital accumulation and economic growth emphasized the two-sided impact of profits, which are both the incentive to invest and also a key source of funds for financing it. She also understood the many obstacles that could prevent an economy from reaching an equilibrium growth path, including resistance by workers to real wage cuts demanded by employers and depressed 'animal spirits' on the part of firms. In her later work, she became a fierce critic of the use of equilibrium methods in economics in general. She emphasized that real-world economies evolve and develop in path dependent ways through historical time, which is not reversible (a point that helps us to understand why protecting domestic industries by imposing tariffs in the present is unlikely to bring back the industries or jobs that left a country in the past).

Robinson's contributions to international economics began with her *Essays in the Theory of Employment* (1947), in which she criticized the use of 'beggar-my-neighbour' remedies for unemployment including import tariffs, industrial subsidies, wage suppression and competitive devaluations. She pioneered in the analysis of how the price and income elasticities of demand and supply of exports and imports affect adjustment in the balance of payments, and how interest rates and investors' expectations affect exchange rates. She was also a trenchant critic of macroeconomic theories of automatic balance-of-payments adjustment, and trade models that assumed balanced trade with full employment. In her later years, she denounced the 'new mercantilism' in which the more advanced countries sought to maintain their advantages over less developed countries, often by advocating 'free trade' for the latter while keeping their own markets closed. In line with her general scepticism about the use of equilibrium models, she was especially critical of the static nature of traditional models of comparative advantage, which she thought ignored the impact of trade on how national economies evolve.

➡

Most fundamentally, those models typically take the resources and technology of countries as given in comparing 'autarky' and 'free trade' equilibria, but in reality Robinson believed that an 'opening up to trade' in historical time would necessarily transform a country's productive capacities rather than merely leading to a more efficient use of those inherited from the past.

EXAM QUESTIONS

True or false questions

1. A country must have an absolute advantage in a good (in the sense of Ricardo) in order to have a comparative advantage in that good.
2. The theory that countries will export only those products in which they have a comparative advantage is valid only if trade is balanced.
3. Free trade typically benefits everyone in a country; no one is ever hurt by trade liberalization.
4. A country can increase production of exports only by diminishing production of other goods, regardless of whether it has full employment or significant unemployment.
5. In the real world, whether a country can export a good (of a given quality) depends on the monetary price it can sell the good at, not whether the country has the lowest opportunity cost in that product.
6. Modern trade agreements (since the 1990s) simply reduce trade barriers, and do not generally limit countries' internal economic policies.
7. Historically, most countries that have reached high-income status (other than oil exporters) did so through industrialization, and often used protectionist policies to achieve that.
8. A developing country may sometimes have to adopt 'infant-industry protection' in order to reap greater gains from trade in the long run, instead of specializing according to its initial comparative advantage.
9. There are no limits to how many manufactured goods developing countries can sell in advanced country markets.
10. It would be easier to obtain political support for free trade (open markets) if governments committed themselves to use macroeconomic policies to maintain full employment.

Multiple choice questions

1. According to the theory of comparative advantage, a country should export the goods for which it has:
   a) the lowest labour cost;
   b) the lowest opportunity cost;
   c) the lowest monetary cost;
   d) the lowest productivity.
2. In the example given in Table 15.2, East Asia has:
   a) the absolute advantage in rice;
   b) the comparative advantage in rice;
   c) the absolute advantage in televisions;
   d) the comparative advantage in televisions.

3. In the example given in Table 15.2, US television workers will:
   a) lose jobs because they are absolutely less productive than East Asian workers;
   b) gain jobs because they are absolutely more productive than East Asian workers;
   c) lose jobs because the USA does not have a comparative advantage in televisions;
   d) gain jobs because the USA has achieved a comparative advantage in televisions.
4. A rise in the terms of trade for the exports of a country that is specialized in primary commodities guarantees that:
   a) the country will capture more of the consumer gains from trade in the short run;
   b) the country will shift into manufactures and achieve rapid growth in the long run;
   c) the country's currency will fall in value, making its other exports more competitive;
   d) the country will be able to escape from boom-bust cycles and financial crises.
5. In a free trade area, each member country must:
   a) not impose tariffs on all (or most)imports from other member countries;
   b) adopt common external tariffs against all non-member countries;
   c) allow free mobility of labour and capital from other member nations;
   d) adopt a common currency and harmonize all economic regulations.
6. Which of the following is required in a customs union?
   a) Labour and capital mobility.
   b) Common external tariffs.
   c) A single currency.
   d) Rules of origin.
7. Some of the reasons why countries seek to join preferential trade agreements may include all of the following, except:
   a) to attract foreign direct investment by multinational corporations;
   b) to obtain preferential market access to large markets such as the USA or EU;
   c) to achieve greater economies of scale for domestic producers;
   d) to get greater gains for domestic consumers through trade diversion.
8. Which of the following is the best example of a global value chain?
   a) Middle Eastern petroleum is sold in European markets.
   b) Mexican automobile parts are used to produce cars in Canada.
   c) French wines compete with California wines in the USA.
   d) Pakistan exports cotton textiles made with Pakistani cotton.
9. According to the Stolper–Samuelson theorem, the winners from international trade liberalization in a given country are likely to be the:
   a) owners of the abundant factor of production;
   b) owners of the scarce factor of production;
   c) owners of all factors of production;
   d) owners of large multinational corporations.
10. Trade agreements can foster increased inequality if they:
   a) increase protections for intellectual property rights of large firms;
   b) create more job opportunities only for the most skilled workers;
   c) allow large multinational corporations to displace local firms;
   d) all of the above.

16

Balance-of-payments constrained growth

John McCombie and Nat Tharnpanich

OVERVIEW

This chapter:

- explains why, for many countries, their long-run rate of economic growth is constrained by their balance of payments to be below their maximum possible growth rate; as such, it presents an alternative, and a more plausible, explanation of growth performance compared with neoclassical growth theory;
- shows that this maximum growth rate of output compatible with balance-of-payments equilibrium is determined by the country's income elasticities of demand for its exports and imports and the growth of world income; this relationship, known as 'Thirlwall's Law', gives a good prediction for the rates of economic growth of a large number of countries;
- shows that non-price competitiveness (as reflected in international differences in the values of both the income elasticities of demand for exports and imports) is more important than price competitiveness in international trade;
- shows that, unlike in the neoclassical approach to the balance of payments, a change in price competitiveness does not significantly affect the rate of economic growth consistent with balance-of-payments (current account) equilibrium;
- demonstrates that it is the growth of output that has to adjust to ensure that the balance of payments is in equilibrium;
- explains how a decline in the growth rate of world income can limit the economic growth of a country.

KEYWORDS

- **Balance-of-payments constrained growth rate:** It is the growth rate of output consistent with the balance of payments (current account) being in equilibrium.
- **Export and import demand functions:** These represent the demand for exports as a function of world income and relative prices, and the demand for imports as a function of domestic income and relative prices.
- **Income elasticity of demand for exports (ε) and imports (π):** These measure the percentage change in the demand for exports divided by the percentage change in world income and the percentage change in the demand for imports divided by the percentage change in domestic income respectively.
- **Non-price competitiveness:** It is competitiveness measured in terms of the quality and sophistication of goods and services and the effectiveness of their distribution networks. It is reflected in the relative size of both the world income elasticity of demand for a country's exports (ε) and the domestic income elasticity of demand for imports (π).
- **Growth of potential output:** It is the maximum possible rate of growth of output and is also called the natural rate of growth.
- **Thirlwall's Law:** It is the equation that gives the balance-of-payments constrained growth rate, which is determined by the ratio of the income elasticities of demand for exports and imports multiplied by the growth of world income.

Why are these topics important?

The growth of productivity (output per head) and output determine the growth of the standard of living in a country. Understanding why these differ between countries is important in an open economy. The heterodox view emphasizes the growth of exports and the role of the balance of payments in limiting the rates of economic growth of some countries to below their maximum values, namely the growth of their potential output (or their natural rate of growth). This contrasts with the mainstream view of economic growth, where the emphasis is not on demand, but on the rate of technical progress and the growth of the factors of production and, in the long run, output grows at its maximum possible value.

The traditional mainstream view

The mainstream view of modelling economic growth within a macroeconomic framework has its origins in a famous paper by Solow (1956). Solow's model is essentially a supply-oriented approach, based on the aggregate production function. This is a technological relationship that specifies

output as determined by its inputs, most notably labour (sometimes adjusted for human capital proxied by the level of education) and capital. The effectiveness with which these factors of production are used is determined by the level of technology. It is assumed that output is always at its full employment level, markets are perfectly competitive and there are constant returns to scale. A surprising result, at the time, was that increasing the rate of investment would not increase the long-run, or steady-state, rate of economic growth. This is because capital is assumed to be subject to diminishing returns. Increasing the capital stock, holding other factors constant, would increase output, but at a diminishing rate. Solow showed that the steady-state growth of productivity (output per head) is determined solely by the exogenously given rate of technical progress.

In this approach, the growth rate of productivity differs between countries because, for some unspecified reason, initially the amount of capital per worker for some countries (or regions) is, for example, below their steady-state levels. These countries therefore experience a temporary acceleration in their growth rates as they catch up. This is termed 'convergence', and there is empirical support for this at the regional level for the United States and Japan (Barro and Sala-i-Martin, 2004, Ch. 11). However, there is no evidence of convergence when all countries are considered together. The model also treats the open economy in a very limited way. It merely allows countries to borrow and lend unlimited amounts on world capital markets at the world interest rate (Williamson, 1983, pp. 113–17).

This led to a second generation of mainstream models that can explain divergence. One way is to assume that the growth of capital is associated with an externality, such as 'learning by doing', which offsets the diminishing returns. Another way is to have a separate production function for the generation of technical change or the production of new ideas. Consequently, the model includes a research and development (R&D) sector. These approaches are often referred to as endogenous growth theory, as the rate of technical change is now explained by the model, and not assumed to be exogenous as in Solow's model.

Consequently, the new mainstream approach can encompass both convergence and, more importantly, divergence in national productivity and output growth rates. But in all the models, factors of production are fully employed and used efficiently. See any macroeconomics textbook and, for a sceptical view, Thirlwall (2002; 2013, Chs 2–3). Trade is considered only to the extent that the degree of openness is often found to be important in determining productivity growth in empirical studies.

The heterodox alternative

By way of contrast, the heterodox approach assumes that the long-run rate of growth of an economy can be below its maximum possible growth rate, or its growth of potential output. It adopts a Keynesian perspective. Unlike the mainstream approach, it emphasizes the importance of the role of the growth of demand, to which supply adjusts, in determining the long-run rate of growth of output. For example, developing countries have high rates of open and disguised unemployment, dual economies and low rates of investment, so growth is not at its maximum rate. Developed countries can also be growing below their growth of potential output even though they have low rates of measured unemployment. An increase in the growth of demand for goods and services leads to an increase in the rate of capital accumulation and expenditure on R&D. Both these can lead to a higher rate of labour-saving technical progress. The use of more capital per worker can save labour and increase productivity. A faster growth of output can induce a faster growth of labour productivity as evidenced by 'Verdoorn's Law' (McCombie et al., 2002). This is an empirical relationship that shows that an increase in manufacturing output growth by, for example, one percentage point increases the growth of productivity by half a percentage point. This is due to increasing returns, broadly defined, and induced technical progress. A faster growth of in-migration can also help offset any labour shortages. Finally, faster economic growth leads to a more efficient use of existing resources and can also lead to the transfer of labour from low- to high-productivity sectors.

Consequently, the crucial question is: what prevents countries always growing in the long run at their maximum possible rate, as the mainstream approach assumes? The answer lies in the role of the balance of payments in potentially restricting the growth rate. As the growth of output increases, so will the growth of imports through the import demand function. Consequently, if import growth exceeds that of exports, this will lead to a growing balance-of-payments (current account) deficit. This has to be financed by the growth of net capital inflows from abroad, which a country cannot sustain for any length of time.

Importantly, it will be shown that the only way to bring the balance of payments into equilibrium is to reduce the rate of growth of output or income. (The terms 'output' and 'income' will be used interchangeably.) This has the result that the economic growth of the country is constrained to be below its maximum possible growth rate. Under these circumstances, the country is said to be growing at its balance-of-payments equilibrium constrained growth rate. Consequently, to achieve a faster growth rate involves either

increasing the growth of exports or reducing the growth of imports, or both. This is necessary to relax the balance-of-payments constraint.

To elaborate this argument, we will consider an extension to the model of balance-of-payments constrained growth, that was first developed by Thirlwall in his seminal paper published in 1979, and still remains the core model.

Determining the balance-of-payments constrained growth rate

The balance-of-payments accounting identity is written as follows:

$$\textit{current account} + \textit{capital account} + \textit{official settlements balance} \equiv 0$$

The current account is the difference in revenues from a country's exports and payments for imports. It also includes the net revenues from interest payments, dividends and wages from overseas, but for simplicity these shall be ignored. The official settlement balance is the change in a country's official foreign currency reserves, which we also shall not consider. The capital account consists of all the changes that alter the external assets and/or liabilities of a country.

Consequently, from the balance-of-payments accounting identity, the current account (CA) may be expressed as follows:

$$CA = P_d X - P_f EM \tag{16.1}$$

where P_d and X are the price in domestic currency and the volume of exports respectively. P_f is the price of imports in foreign currency and E is the nominal exchange rate, measured as the domestic price of foreign currency. M is the volume of imports.

If the current account is in deficit (that is, the value of the expenditure on imports exceeds the revenues from exports), this difference has to be financed by capital flows from abroad. We may therefore express the balance-of-payments identity as $P_d X + F \equiv P_f EM$; that is, the total revenues from exports plus the value of net capital inflows, measured in the domestic currency (F), is equal to the expenditure on imports. The balance-of-payments constraint arises because, like an individual, a country cannot go on borrowing indefinitely.

Two indicators widely used to measure a country's degree of foreign indebtedness are the ratio of the current account deficit to gross domestic product

(GDP), and the net stock of foreign liabilities (NFL) to GDP. The NFL is the total net amount that a country owes to the rest of the world. Empirical evidence suggests that for both advanced and developing countries, the probability of a currency crisis increases dramatically once the NFL exceeds 50 per cent of GDP or is about 20 percentage points greater than its historic average. Moreover, the speed at which the NFL increases, which depends on the size of the current account deficit, is also a warning indicator (Catão and Milesi-Ferretti, 2014). However, foreign direct investment, where an investor or company in one country directly invests in an enterprise in another country, poses fewer problems. This is because the investment is long-term and normally directly adds to the productivity of the recipient country and the capital usually does not have to be repaid. Remittances from a country's nationals working abroad can also be an important source of foreign exchange, especially for some developing countries.

A balance-of-payments crisis occurs when the foreign lenders become concerned that the current account deficit will lead to a sharp fall in the exchange rate. This will cause a sudden steep fall in the value of their loans in terms of the lenders' domestic currency. In other words, it will cause them a large capital loss. There may also be the fear of default on any of their loans, even those made to the government. All this will lead to an unwillingness to roll over the loans, or worse, to a sudden capital flight. The value of the portfolios of shares that foreign investors hold in the country will also fall in terms of the investors' domestic currency. This may lead them rapidly to sell the shares and convert the proceeds into their domestic currency. A sudden crisis then occurs and the current account is brought back into equilibrium by a fall in the level of output and a reduction in output growth. This may be done by, for example, economic policy deflating the economy to save foreign exchange and raising interest rates in order to prevent capital flight. Private investment may also fall as entrepreneurs' expectations about the economic outlook worsen. Eichengreen and Gupta (2016) have identified 46 'sudden stops' in capital inflows leading to a fall in GDP growth in 34 emerging countries between 1991 and 2015. However, it should be noted that a country can be balance-of-payments constrained without a crisis actually occurring.

For expositional ease, let us assume to begin with that the balance-of-payments equilibrium rate of growth occurs when there is no growth in net capital flows. In this case, the growth of the balance of payments, starting from equilibrium, is derived from equation (16.1) and is given by the following equation:

$$p_d + x = p_f + e + m \tag{16.2}$$

where the lower case of a variable denotes its growth rate. To see how this acts as a constraint on economic growth, it is necessary to consider next what determines the growth of exports and imports. In this Keynesian approach, the growth of a country's exports is determined by the growth of demand for them. This is determined by the growth of the income of the countries with which the country trades, weighted by their trade shares, namely w, together with the country's rate of change of international price competitiveness. (For simplicity, we term the variable w the growth of 'world' income.) The growth of demand for exports may be expressed as follows:

$$x = \varepsilon w + \eta(p_d - p_f - e) \tag{16.3}$$

where ε is the world income elasticity of demand for exports and measures how much the growth of demand for exports increases as a result of the increase in world income. If ε takes a value of, say, 2 then an increase in the growth of world income by one percentage point increases the growth of a country's exports by two percentage points. η, which takes a negative value, is the price elasticity of the demand for exports. Thus, the growth of the country's exports is determined by the growth of world income and the rate of change of the country's relative prices in terms of its domestic currency. $(p_d - p_f - e) < 0$ represents the rate of change of the improvement in the country's relative price competitiveness.

The growth of imports is given by the following equation:

$$m = \pi y + \psi(p_f + e - p_d) \tag{16.4}$$

where π is the income elasticity of demand for imports, y is the growth of the country's income and ψ (which is negative) is the price elasticity of demand for imports. Equation (16.4) shows that the growth of a country's income determines the growth of its imports, and the latter is also affected by the rate of change of its relative prices.

If we substitute equations (16.3) and (16.4) into (16.2), the equation for a country's balance-of-payments equilibrium growth rate is derived as follows:

$$y_B = \frac{1}{\pi}(1 + \eta + \psi)(p_d - p_f - e) + \frac{\varepsilon}{\pi} w \tag{16.5}$$

In other words, y_B is the hypothetical fastest rate of economic growth that the country can achieve compatible with its balance of payments (or current account) being in equilibrium. If this is below the maximum possible growth rate that the country could achieve (that is, below its growth rate of potential

output or y_P), then economic growth is said to be balance-of-payments constrained.

The first term in equation (16.5) is the contribution to y_B of the rate of change of the country's price competitiveness and the second term is that of the growth of world income.

The values of the various parameters in the model are obtained from the statistical estimation of the export and import demand functions. Empirically, the rate of change of relative prices is found to play only a very small role in determining the growth of imports and exports. For the rate of improvement of price competitiveness to increase y_B (that is, when $(p_d - p_f - e) < 0$), it is necessary that $(\eta + \psi) < -1$, which means that $(1 + \eta + \psi)$ is negative. This is known as the Marshall–Lerner condition, which is necessary for an improvement in the balance of payments. If the sum of $\eta + \psi$ equals -1, and so the Marshall–Lerner condition is not met, it can be seen from equation (16.5) that the rate of change of relative prices will have no impact on the balance-of-payments constrained growth rate. Many econometric estimations of the export and import demand functions of different countries find that usually the Marshall–Lerner condition is either not, or only just, met. See, for example, the extensive survey by Bahmani et al. (2013).

Moreover, it can be seen from equation (16.5) that to increase the growth of exports (or to reduce the growth of imports) there needs to be a continuous fall in relative prices. However, it is difficult to generate this by a continuous depreciation of the nominal exchange rate. This is because, for example, a one-off depreciation of the exchange rate increases the prices of imports in terms of the domestic currency. If this is factored into wage bargaining in an attempt to prevent a fall in real wages, then eventually the consequent rise in domestic prices will partially, or fully, offset the effect of the initial depreciation. This is termed the 'real wage resistance' effect. Real wage resistance will also reduce the effect of a continuous depreciation of the exchange rate as the resulting increase in the growth of nominal wages leads to an increase in the growth of domestic prices. Moreover, many imports are inputs into domestically produced goods and services and this will also increase domestic prices.

Finally, in the face of a depreciation of the nominal exchange rate, some foreign suppliers of imports to the country may choose to try to ensure that the prices of their goods and services in terms of the country's currency remain competitive. They do this in order not to lose market share in the country. A depreciation of the exchange rate will increase the foreign

suppliers' prices in terms of the country's currency. To stay price-competitive with import-competing domestic firms, the foreign suppliers need to reduce the price of their exports P_f to offset the effects of the depreciation, and this will be at the expense of their profits. This is known as 'pricing to market'. In growth rate terms, the effect can be seen from equation (16.4). This is the country's import function, which also determines the growth of the exports of its foreign suppliers. If e initially increases, then a reduction in p_f by the foreign suppliers will have the effect of offsetting, to a greater or lesser extent, the effect of the increase in e. The effect of real wage resistance and pricing to market are likely to go a long way in preventing relative prices, and their growth rates, from greatly changing.

Consequently, over the medium to long term, $(p_d - p_f - e)$ does not substantially change. This does not mean that there are no sudden and substantial shocks to the rate of change of relative prices, but these are usually only temporary. See for example the case study of the 1997 Asian financial crisis discussed in Box 16.1.

It is important not to confuse the generally small variation in the growth of relative prices, in the medium to long run, with the neoclassical 'law of one price'. The latter occurs where homogeneous goods are traded and a country can export as much as it wants. This is because the price elasticity of demand for exports is, strictly speaking, minus infinity and so the growth of relative prices only has to change by a very small amount to ensure that the balance of payments is in equilibrium. Under this implausible assumption, economic growth can never be constrained by the balance of payments. (The neoclassical monetary approach to the balance of payments, for example, makes this assumption.)

Given the above discussion, with the rate of change of relative prices having little or no effect, the balance-of-payments constrained growth rate is given by the second term in equation (16.5), namely:

$$y_B = \frac{\varepsilon}{\pi} w \tag{16.6}$$

As $x = \varepsilon w$ and if $(p_d - p_f - e) = 0$, this may be expressed as follows:

$$y_B^* = \frac{1}{\pi} x \tag{16.7}$$

These two equations have come to be known as 'Thirlwall's Law'.[1] Consequently, the ratio of the income elasticities of demand for exports and

BOX 16.1

THE ASIAN FINANCIAL CRISIS OF 1997

The Asian financial crisis occurred in July 1997, when the Thailand baht, which had previously been pegged to the US dollar, went into freefall. There was a contagion effect and the currencies of Indonesia, Malaysia, the Philippines and South Korea also fell sharply. Output in these countries experienced a sudden collapse, which had a knock-on effect for the world economy.

The crisis caught most policymakers (including the International Monetary Fund) by complete surprise. Up until then, these countries had been experiencing exceptionally fast growth rates. The only possible problem was the large balance-of-payments deficits that especially Thailand and Indonesia had been running. In the case of Thailand, this was financed by short-term volatile capital flows that went largely as speculative loans to Thai real estate companies. This led to real estate and asset bubbles that attracted further foreign capital. The triggering of the crisis occurred because the Thai government did not have the financial resources to bail out a failing large domestic bank. A consequence was a sudden massive capital flight out of Thailand, and collapse in the exchange rate. This sent out a warning signal to foreign lenders to the other Asian countries, who also rapidly withdrew their capital loans. If they had not done this, the capital loss they would have experienced as the exchange rate fell would have substantially outweighed any gains from the higher interest rates and returns in these countries.

The impact on these economies was severe. Many loans were denominated in US dollars and the cost of paying these off in domestic currency rocketed to such an extent that many Asian firms became technically insolvent. This led to a rise in non-performing loans and there was a severe credit squeeze. The resulting steep fall in investment and output led to a collapse in imports. It was this that led, in turn, to a subsequent improvement in the balance of payments and a subsequent appreciation of the exchange rates. The crisis provided a salutary reminder of the important role of the balance of payments in constraining the growth of countries.

imports, that is, ε/π, becomes important in explaining why the balance-of-payments constrained growth rates differ between countries.

Equations (16.6) and (16.7) are sometimes known as the 'strong' and 'weak' versions of Thirlwall's Law, respectively. The value of ε in equation (16.6) is derived from a regression that explicitly takes account of the change in relative prices, while the use of x in equation (16.7) assumes that its effect is negligible.

Consider, for example, Italy in the early post-war period, 1951–73. Its estimated value of ε was equal to 2.93 and that of π was 2.25. With the growth of world income at about 4 per cent per annum, y_B took a value of 5.21 per

cent per annum. (Italy's actual growth rate over this period was 5.1 per cent per annum.)

Compare this with the United Kingdom (UK) over the same period, when ε took a value of 1.12 and π was 1.51. Consequently, y_B was 2.97 per cent per annum, compared with its actual growth rate of 2.7 per cent. This was the time of the 'Golden Age' of rapid growth of the industrialized countries. At the time, the UK government saw no reason why the United Kingdom should not be able to match the faster growth rates of continental Europe. But as UK growth accelerated, imports grew much faster than exports, leading thereby to substantial current account deficits. Consequently, the UK's growth has been described as 'stop-go' over this period, and there were major balance-of-payments crises in 1967 and 1976. In the last case, the International Monetary Fund had to be called in to bail out the pound sterling.

The crucial question, therefore, is: what determines the values of ε and π and why do they differ between countries? The answer is that the disparities between countries in these values reflect all aspects of non-price competitiveness (see McCombie and Thirlwall, 1994, Ch. 4). As we have already seen, the effect of the rate of change of price competitiveness is negligible. Non-price competitiveness includes the quality and sophistication of exports (and imports), their reliability and the effectiveness of the country's distribution system. Trade is now dominated by intra-industry trade, where countries trade similar goods, partly because consumers like variety and there is product differentiation. An implication is that successful firms compete by introducing better-quality goods or services, which has the effect of shifting their demand curve outwards, rather than by price-cutting and moving down the demand curve.

Non-price competitiveness is also determined by the composition of a country's exports. In other words, some countries will be exporting a combination of categories of goods and services for which world demand is growing relatively fast, such as electronic goods or financial services. Others will be concentrating on sectors, such as basic textiles, for which the growth of world demand is relatively slow. A more sophisticated development of Thirlwall's Law is its multi-sector version. This starts with the fact that the aggregate income elasticity of, for example, exports is the weighted sum of the individual sectoral export elasticities, which will vary in their values. In the course of development, structural change occurs and a successful country is likely to export a larger proportion of goods and services with high income elasticities of demand for exports. This will have the effect of increasing the overall value

of ε, which will increase the growth of total exports for any given w (Araujo and Lima, 2007). Likewise, changes in the sectoral composition of imports will affect the aggregate income elasticity of demand for imports. The importance of this effect has been confirmed by Cimoli et al. (2009) and Gouvea and Lima (2010).

Including the growth of capital flows

The growth of capital flows can be straightforwardly included in the model given by equation (16.6) and implicitly by equation (16.7). With the growth of capital flows, equation (16.2) now becomes $\theta p_d + \theta x + (1 - \theta)f = p_f + e + m$, where f is the growth of nominal net capital inflows measured in terms of the domestic currency. θ is the proportion of export receipts to import expenditure (that is, $P_d X/P_f EM$). If we assume that the rate of change of relative prices is unimportant in determining y_B, then the growth rate consistent with the 'overall balance of payments', following Thirlwall's (2011) terminology, is given by the following equation:

$$y_{OB} = \frac{\theta\varepsilon}{\pi} w + \frac{(1-\theta)}{\pi}(f - p_d) \tag{16.8}$$

The expression $(f - p_d)$ is the growth of real net capital inflows. θ is normally large, say, exceeding 0.9. If the growth of demand exceeds the balance-of-payments equilibrium growth rate, then we can see from the above equation that this will necessitate a growth of net real capital inflows that is likely to be unsustainable in the long run. Hence, economic growth will be equal to the balance-of-payments equilibrium growth rate.

However, even if the growth of real net capital inflows can be sustained if it consists of, say, foreign direct investment, the overall effect on the balance-of-payments constrained growth rate is likely to be quantitatively small. For example, if the growth of real capital inflows is the same as the growth of exports (that is, εw or x), as may be seen from equation (16.8), there is no effect on y_B, which equals y_{OB}. Nevertheless, with the growth of real net capital inflows, overseas lenders have increasing financial claims on the country concerned.

Let us discuss the role of the growth of capital flows further, using a diagrammatical approach.

Balance-of-payments constrained growth: a diagrammatical approach

It is important to note that not all countries will necessarily be balance-of-payments constrained. Some economies, such as those of Japan in the early post-war period or China in the last two decades or so, have had, or are having, such a fast rate of growth of exports that their growth is constrained by the supply side. This may be due to, for example, limits to the rate of capital accumulation and to the rate of the interregional and intersectoral transfer of labour. In these circumstances, such countries can run persistent current account surpluses.

Other countries are what may be termed 'policy-constrained', such as some industrialized countries in the 1970s, where an indirect effect of deflationary policies to curb the rampant inflation was to restrict their rate of growth. More recently, the growth of countries following the 'economics of austerity' in an attempt to reduce budget deficits and the level of government debt may also be policy-constrained.

Consequently, let us denote the maximum possible growth of output by y_P, the actual growth rate of output (supply) by y_S and the growth of demand by y_D. The last of these is determined in the Keynesian approach by the growth of demand, including that of autonomous investment, government expenditure and exports. The relationships between the growth rates of exports, imports, and y_B, y_D and y_P are depicted in Figure 16.1.

Figure 16.1 The growth of exports and imports and the balance-of-payments constraint

Figure 16.1 shows the relationship between a country's growth of imports, exports and income. Import growth, which is given by the line *m*, increases with the growth of domestic income. This is because of the effect of the import demand function (that is, $m = \pi y$). The slope of the line *m* in the figure is π. The growth of exports is determined by the growth of world income (that is, $x = \varepsilon w$) and so does not vary with *y*. Hence the export growth line is horizontal and is given by *x*.

The balance-of-payments constrained growth rate is where the growth of exports equals the growth of imports and is given by the point *A*. At this point $x = m = \pi y_B$ and $y_B = y_B^* = x/\pi$. In equilibrium growth, this will also be equal to the growth of demand, y_D, which in turn determines the growth of supply. As y_B, at point *A*, is below the growth of potential output (that is, $y_B < y_P$), the country is balance-of-payments constrained.

Suppose the country increases its rate of growth of demand (y_D) in an attempt to grow persistently at its maximum rate (y_P). It can be seen that the growth of imports now exceeds the growth of exports and this difference in growth rates is given by *BC*. It will be necessary to finance this by the growth of net capital inflows. But the country will be unable to do this for any length of time, for the reasons discussed above, and so y_D will eventually have to return to y_B. To increase permanently its rate of economic growth, the country needs to increase the non-price competitiveness of its production, which would result in an increase in the ratio ε/π.

In terms of Figure 16.2, an increase in ε will shift the export growth line, *x*, upwards to *x'*. At the same time, if π declines, the import growth line, *m*, will rotate downwards to the right and is given by the new line *m'*. If the combined effect is such that *x* and *m* now intersect at point *D* and the growth of demand increases to y_{1D}, the country will then be growing at its rate of growth of potential output, y_p, and will not be balance-of-payments constrained.[2] This is because the balance-of-payments equilibrium growth rate has increased to y_{1B}. However, increasing non-price competitiveness is difficult and is only likely to occur over a long time period, if at all. Other countries will also be striving to increase their own relative non-price competitiveness. This analysis, nevertheless, shows the importance of export-led growth. This is because the increase in the growth of exports is the only component of production and demand that increases the rate of growth of output and simultaneously earns the foreign currency to pay for the resulting induced increase in import growth.

The nature of the balance-of-payments equilibrium growth rate is shown alternatively in Figure 16.3, which gives a different perspective of the model.

Figure 16.2 The growth of exports and imports and relaxing the balance-of-payments constraint

Figure 16.3 The balance-of-payments equilibrium growth rate and non-price competitiveness

The vertical axis is the growth rate of the country, while the horizontal axis is the world growth rate. The line y_B shows the balance-of-payments equilibrium growth rate for any given w and its slope is given by ε/π. This reflects Thirlwall's Law, namely that $y_B = (\varepsilon/\pi)w$. Any point above the line y_B represents a growing balance-of-payments deficit, and any point below represents a growing surplus.

The line y_D gives the growth of demand as the growth of world income increases. The growth of demand consists of two components. The first is the growth of income due to the growth of the autonomous internal components of demand, such as investment and government expenditure, together with the effect of their dynamic multipliers. This is given by the distance between b and the origin, o, in Figure 16.3. The second component is the growth of income due to the growth of exports and its associated dynamic multiplier. This is given by the distance between y_1 and b. The line y_D slopes upwards because as the growth of world income increases, so does the growth of exports. Consequently, the growth of demand increases with an increase in w, even if the growth of the internal components of demand remains constant. As output growth is demand led, the line y_D also represents the growth of supply, y_S. With world income growing at w_1, point a gives the balance-of-payments equilibrium growth rate and the associated growth rates of income and of demand, that is, y_1.

Suppose that this is below the growth of productive potential and the country attempts to grow at a target rate of y_2. It does this by, say, increasing the growth rates of the internal components of demand through expansionary macroeconomic policy measures. There is no increase in the rate of growth of demand from exports, as the growth of world income does not change. The effect is to shift the line y_D upwards until it passes through the point d and is given by the line y'_D. The increase in the growth of output from the increase in the growth of the internal components of demand is given by cb.

The increase in the rate of growth of output is now greater than y_B, which equals y_1, by the amount da. This will need to be covered by the growth of net capital inflows. This is because the growth of imports now exceeds the growth of exports. However, as we have seen, this situation will be unsustainable, except possibly in the very short run. The growth of the internal components of demand will have to fall until the growth of domestic demand returns to equal y_1. In a sense, the growth of the internal components of demand are not strictly autonomous, as they are dependent on y_B and, hence, indirectly on the growth of exports.

As shown above, the only way that the country can increase its sustainable growth rate is to improve its non-price competitiveness. This increases the value of ε/π, which has the effect of rotating the y_B line upwards to give the line y'_B.

If the country is successful in doing this, the balance-of-payments equilibrium growth rate is given by y_2 at the point *d*. The growth of demand increases and the line y_D shifts upwards until it passes through point *d*, in a similar manner to that discussed above.[3] However, as we have emphasized, increasing the non-competitiveness of a country is not something that can be easily accomplished.

Let us consider the effect of an increase in the growth of world income. This is shown in Figure 16.4. Suppose the country is again growing at its balance-of-payments equilibrium growth rate given by y_1 and the growth rate of world income increases. This has the effect of shifting the w_1 line to the right, to w_2. The balance-of-payments equilibrium growth rate will increase to the rate given by where the w_2 line intersects the line y_B at the point *f*. The growth of output will be given by this point, as the line y_D shifts upwards to y''_D and will also intersect y_B and w_2 at *f*.

Figure 16.4 The balance-of-payments equilibrium growth rate and a change in the growth of world income

Let us assume now that y_2 is the growth of productive potential y_p. If the country is growing at this rate, a faster growth of world income greater than w_2 will not increase y_B. This is because the growth of exports will not increase as they are constrained by the growth of supply, and the growth of imports will remain the same because the growth of income remains constant at y_2 (and at y_P). Consequently, the y_B line becomes horizontal to the right of the point *f*. Likewise an increase in the growth of demand can no longer lead to an increase in the growth rate of output and income, and so y''_D also becomes horizontal at the point *f*.

Figure 16.4 can also be used to show the impact of a world recession and a consequent slowdown in world income. Suppose world income is initially growing at w_2 and this falls to w_1. This means that the country can no longer grow at the rate y_2, because it would now be running an increasing balance-of-payments deficit when the growth of world income is w_1.

Consequently, the growth of demand falls from y''_D to y_D. This may be partly due to a fall in investment expenditure resulting from the fall in export growth and a general worsening of the expectations of entrepreneurs and consumers about the general growth prospects. The economy moves from point *f* to *a* and its growth rate falls to y_1. In this analysis we are looking at the case of only one country, but other countries will also find their growth rates of income falling because of the deterioration of their balance of payments. This will contribute to the fall of w_2 to w_1.

An implication is that, in these circumstances, no individual country by itself can increase its growth rate by, say, macroeconomic policies. As explained above, if it attempts to do so, it will run into balance-of-payments problems. What is needed is a coordinated expansion by all countries. This will increase the growth of world income back towards w_2. This was the policy adopted, for example, at a summit meeting of a group of leading industrialized and emerging countries (the Group of 20 or G20) in London in 2009, during the severe collapse in world output as a consequence of the subprime financial crisis. The G20 committed themselves to a substantial coordinated fiscal expansion of some US$5 trillion and to make available an extra US$1.1 trillion to help the International Monetary Fund and other global institutions to boost growth and employment. Some Keynesian lessons seem to have been learned.

Tests of the model and empirical evidence

Since the publication of Thirlwall's (1979) seminal paper, there have been numerous studies that have both extended and tested the balance-of-payments constrained growth model. In essence, these tests of the model are to see how closely the estimate of y_B predicts the actual growth rate (y_A) of countries (McCombie and Thirlwall, 1994). y_B is calculated as $\hat{\varepsilon}w/\hat{\pi}$, or $x/\hat{\pi}$, where $\hat{\varepsilon}$ and $\hat{\pi}$ are the econometric estimates of the elasticities using time-series data and y_B, w and x are the average growth rates, calculated over a period of several years. The estimation techniques have progressively become more sophisticated and recent estimation techniques have allowed the estimates of the long-run elasticities to change over time. It is generally found that the estimates of the income elasticities are statistically significant and those of the relative price variables are not. Consequently, when the average growth of y_B and the average actual growth rate y_A are compared, they are found to be close. As Thirlwall (2011, p. 25) puts it, 'if relative prices were an efficient balance of payments adjustment mechanism, no necessary relation between the two rates is to be expected'. Moreover, if the estimates of ε and π are found to be statistically insignificant, this would be sufficient to reject the law, but this is rarely the case.

There have been numerous tests used to see whether the difference between y_A and y_B is statistically significant (see Thirlwall, 2011; 2013, pp. 111–18). Generally, ever since Thirlwall (1979) found a close correspondence between y_B and y_A for the advanced countries for the period 1951–73, it is found that they are not greatly different. (See Thirlwall, 2011, for a discussion of the various methods of testing the hypothesis.)

Differences in the relative income elasticities of demand for imports and exports are only proxies for disparities in non-price competitiveness. Additional variables to capture the degree of non-price competitiveness such as R&D expenditures have also been included in the import and export demand functions.

The interpretation of Thirlwall's Law has not been without a few erroneous criticisms. See, for example, the discussion in McCombie and Thirlwall (1994). A recent criticism repeats the argument that Thirlwall's Law is merely an identity, a critique that was shown to be wrong nearly 50 years ago. It confuses the statistical estimate of an elasticity with one that calculates it arithmetically. For further details, see McCombie (2019).

The importance of non-price competitiveness in international trade has been emphasized by Benkovskis and Wörz (2015) in their aptly titled paper – 'Cost

competitiveness: a dangerous obsession'. They ask: 'How can we reconcile real effective exchange rate (REER) – and hence an apparent deterioration in cost competitiveness – with rising world market share – a clear sign of improved competitiveness?' (Benkovskis and Wörz, 2015). (This is also known as the Kaldor paradox; Kaldor, 1978.) Benkovskis and Wörz point out that many developed and emerging market economies have increased their share of world markets although they have increasing relative prices, while other countries have lost global market share despite increasing cost-competitiveness. They show quantitatively that this is due to changes in non-price competitiveness that have outweighed any changes in relative prices.

Concluding comments

Thirlwall's Law is a parsimonious explanation of why rates of economic growth differ, but the mark of a good economic model is to explain a lot from a little. This chapter has discussed the simplest version of the balance-of-payments constrained equilibrium growth model and where world growth is taken to be exogenous. However, the model can be generalized to the case of two or more trading blocs or groups of countries. It also provides an explanation of why regional problems are balance-of-payments problems (Thirlwall, 1980). Moreover, recently the importance of non-price competitiveness, compared with price competitiveness, has increasingly become recognized by economists, although it has always been taken for granted in business schools. While the balance-of-payments constrained growth model is highly aggregated, it is not more so than the mainstream growth models. However, a major difference is that the latter treat output as homogeneous, in which the structure of production and trade does not matter, whereas the assumption about the qualitative differences between countries' exports (and imports) is crucial in Thirlwall's model. It shows why the long-run growth of a country in an open economy may be below its maximum growth rate, and emphasizes the importance of the growth of demand in a long-run context, as well as in the short run.

NOTES

1 Under certain assumptions, this is the dynamic version of Harrod's foreign trade multiplier (Harrod, 1933). See, for example, Thirlwall (2002, p. 71).

2 It should be noted that it is possible for the m line to rotate upwards if π increases, so long as ε increases sufficiently for x' and m' still to intersect on the y_p line. The requirement is for the ratio ε/π to increase sufficiently to achieve this. However, a fall in π will occur if the non-price competitiveness of the country's import-competing industries increases, as assumed in Figure 16.2.

3 A minor difference, however, compared with the previous case, is that the new growth of the demand curve (not shown in Figure 16.3) is slightly steeper than the line y'_D and the contribution of the autonomous components of demand is correspondingly less. This is because the growth of exports, due to the increase in non-price competitiveness, and its contribution to the resulting growth in demand are now greater than in the previous example.

REFERENCES

Araujo, R.A. and G.T. Lima (2007), 'A structural economic dynamic approach to balance-of-payments-constrained growth', *Cambridge Journal of Economics*, **31** (5), 755–74.

Bahmani, M., H. Harvey and S.W. Hegerty (2013), 'Empirical tests of the Marshall–Lerner condition: a literature review', *Journal of Economic Studies*, **40** (3), 411–43.

Barro, R.J. and X. Sala-i-Martin (2004), *Economic Growth*, 2nd edition, Cambridge, MA: MIT Press.

Benkovskis, K. and J. Wörz (2015), 'Cost competitiveness: a dangerous obsession', *VOX CEPR Policy Portal*, accessed 18 December 2019 at http://voxeu.org/article/cost-competitiveness-obsession.

Catão, L.A.V. and G.M. Milesi-Ferretti (2014), 'External liabilities and crises', *Journal of International Economics*, **94** (7), 18–32.

Cimoli, M., G. Porcile and S. Rovira (2009), 'Structural change and the BOP-constraint: why did Latin America fail to converge?', *Cambridge Journal of Economics*, **34** (2), 389–411.

Eichengreen, B. and P. Gupta (2016), 'Managing sudden stops', World Bank Policy Research Working Paper, No. 7639.

Gouvea, R.R. and G.T. Lima (2010), 'Structural change, balance-of-payments constraint, and economic growth: evidence from the multisectoral Thirlwall's law', *Journal of Post Keynesian Economics*, **33** (1), 169–204.

Harrod, R. F. (1933), *International Economics*, Cambridge, UK: Cambridge University Press.

Kaldor, N. (1978), 'The effect of devaluations on trade in manufactures', in N. Kaldor (ed.), *Further Essays on Applied Economics*, London: Duckworth, pp. 99–116.

McCombie, J.S.L. (2019), 'Why Thirlwall's Law is not a tautology: more on the debate over the law', *Review of Keynesian Economics*, **7** (4), 429–43.

McCombie, J.S.L., M. Pugno and B. Soro (eds) (2002), *Productivity Growth and Economic Performance: Essays on Verdoorn's Law*, Basingstoke: Macmillan.

McCombie, J.S.L. and A.P. Thirlwall (1994), *Economic Growth and the Balance-of-Payments Constraint*, Basingstoke: Macmillan.

Solow, R.M. (1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70** (1), 65–94.

Thirlwall, A.P. (1979), 'The balance of payments constraint as an explanation of international growth rate differences', *Banca Nazionale del Lavoro Quarterly Review*, **32** (128), 45–53.

Thirlwall, A.P. (1980), 'Regional problems are "balance-of-payments" problems', *Regional Studies*, **14** (5), 419–25.

Thirlwall, A.P. (1987), *Nicholas Kaldor, Grand Masters in Economics*, Brighton, UK: Wheatsheaf.

Thirlwall, A.P. (2000), *The Euro and 'Regional' Divergence in Europe*, London: New Europe Research Trust.

Thirlwall, A.P. (2002), *The Nature of Economic Growth: An Alternative Framework for Understanding the Performance of Nations*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Thirlwall, A.P. (2011), 'Balance of payments constrained growth models: history and overview', *PSL Quarterly Review*, **64** (259), 307–51.

Thirlwall, A.P. (2013), *Economic Growth in an Open Developing Economy: The Role of Structure and Demand*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Thirlwall, A.P. (2018), 'A life in economics', *PSL Quarterly Review*, **71** (284), 9–39.

Thirlwall, A.P. and P. Pancheco-López (2017), *Economics of Development: Theory and Evidence*, 10th edition, London, UK and New York, USA: Palgrave Macmillan.

Williamson, J. (1983), *The Open Economy and the World Economy*, New York: Basic Books.

A PORTRAIT OF ANTHONY PHILIP THIRLWALL (1941–)

Anthony Philip Thirlwall is a British macroeconomist who has made distinguished contributions to economics in showing that the insights of Keynes, and the subsequent developments of Keynes's theory, are still very important for the understanding of the modern macroeconomy. This is notwithstanding the dominance now of the mainstream (neoclassical) macroeconomic approach that largely discounts the relevance of the Keynesian approach. For example, the mainstream approach assumes that the macroeconomy is self-stabilizing, apart from some minor economic frictions due to sticky prices. The Keynesian approach shows that unemployment can persist because of the lack of demand in the product market, which has nothing to do with rigid prices. As Thirlwall (2018) writes in his autobiographical article, he is proud to call himself an unreconstructed Keynesian. Thirlwall (2018, p. 13) states that 'there are few areas of mainstream macroeconomics that cannot be satisfactorily analysed with the concepts and insights that Keynes bequeathed', a sentiment shared by many heterodox economists. He has presented compelling criticisms of the 'natural rate of unemployment' and of the 'natural rate of growth', concepts central to orthodox macroeconomics. He helped to keep the Keynesian flame burning at the University of Kent, where he has spent nearly all of his academic career, including the organization of 11 influential biennial Keynes seminars.

His contributions to macroeconomics are wide-ranging, including important criticisms of the orthodox approach to macroeconomics, regional economics and development economics. Thirlwall's early work was in regional economics and during this time he formalized (with Robert Dixon) a regional cumulative causation model of economic growth. This had been outlined by Nicholas Kaldor in verbal form. Kaldor approved of this approach, and Thirlwall found his own approach to economics closely mirroring that of Kaldor. One result of this was that he wrote Kaldor's intellectual biography (Thirlwall, 1987). Thirlwall further formalized and developed some of the other rich insights of Kaldor into the workings of the macroeconomy. The balance-of-payments constrained model discussed in this chapter (Thirlwall, 1979) arose out of Thirlwall's early interest in the UK's balance-of-payments problems. While there have been developments and extensions of this approach, the 1979 paper is still the canonical model. He further showed that regional problems are in fact balance-of-payments problems, in spite of regions having a single currency (Thirlwall, 1980). This raises serious problems for the growth of eurozone countries, which is now affecting the peripheral EU countries, as Thirlwall (2000) presciently pointed out some years ago.

Much of his research has been in the area of development economics, where he has made many notable contributions. He not surprisingly views development from a Keynesian–Kaldorian perspective. In particular, his more recent research (with Penélope Pacheco-López) has questioned the conclusion that trade liberalization is always beneficial to the liberalizing country.

➡

The reduction in tariffs can often lead to the growth of imports increasing faster than exports, with attendant balance-of-payments problems.

He has written (now co-authored with Pacheco-López) a best-selling textbook, recently retitled *Economics of Development*. It was first published in 1972 and is presently in its tenth edition, published in 2017. Although Tony Thirlwall has now retired, he still continues to be highly productive in research and he was, for many years, the General Editor of The Great Thinkers in Economics series published by Palgrave Macmillan. His research has helped to keep the Keynesian approach alive and well, in spite of mainstream economics moving in the opposite direction.

EXAM QUESTIONS

True or false questions

1. In neoclassical growth theory, there are no demand constraints on the rate of output growth.
2. The main determinant of a country's export growth is its growth of output.
3. The balance-of-payments constrained growth model implies that countries can grow persistently below their growth of productive potential.
4. When a country is balance-of-payments constrained, the ultimate determinant of the growth of output is the growth of investment.
5. The income elasticity of demand for exports measures the growth of exports with respect to the growth of its relative price.
6. Regions within a country cannot be balance-of-payments constrained if they trade only with each other.
7. Estimates of the UK's price elasticity of demand for exports and imports are –0.62 and –0.32, respectively. This implies that a rate of decline of relative prices will not raise its balance-of-payments equilibrium growth rate.
8. A successful strategy for all countries to raise their balance-of-payments equilibrium growth rate is simultaneously to impose import quotas that will reduce their income elasticity of demand for imports.
9. The growth in intra-industry trade (for instance, Germany and the United Kingdom export cars to each other) reflects the increasing significance of variety as a cause of trade and, hence, the growing importance of non-price competitiveness.
10. Thirlwall's Law states that a country's growth rate can be predicted by the ratio of its growth of exports to its growth of imports.

Multiple choice questions

1. Balance-of-payments equilibrium on the current account means that net capital flows are:
   a) negative;
   b) positive;
   c) zero;
   d) any of these.
2. In the balance-of-payments constrained growth model, which variable adjusts to bring the current account into long-run equilibrium?

a) The growth of income.
b) The growth of capital flows.
c) The growth of relative prices.
d) The growth of exports.

3. Over the period 1951–73, Japan's growth of exports was 15.4 per cent per year and its income elasticity of demand for imports was 1.23. Calculate its balance-of-payments equilibrium growth rate. If its actual growth rate was 9.5 per cent per annum, this means that the growth of Japan was:
a) balance-of-payments constrained by export growth;
b) balance-of-payments constrained by import growth;
c) constrained by the growth of net positive capital flows;
d) supply, or policy, constrained.

4. If the nominal exchange rate depreciates by 4 per cent per year, domestic prices increase by 5 per cent per year, and foreign prices grow at 2 per cent per year, the rate of change of relative prices in the balance-of-payments constrained growth model will be:
a) –3 per cent per year;
b) –1 per cent per year;
c) 1 per cent per year;
d) 2 per cent per year.

5. If a country's nominal exchange rate is depreciating and the foreign suppliers of its imports set their prices in the country's currency to maintain their market share (that is, they price to market) this will:
a) decrease profits;
b) increase profits;
c) make no difference to their profits;
d) any of these.

6. There is initially no growth in the price of exports and imports measured in the domestic currency. The price of exports now declines by 3 per cent per annum and the growth of world income remains the same. The price elasticity of demand for exports is –0.5. Using the export demand function expressed in growth rate form, the percentage point change in the growth in volume of exports (x) and in export revenues ($p_d + x$) will be, respectively:
a) 1.5 and 1.5;
b) 1.5 and –1.5;
c) –1.5 and 1.5;
d) –1.5 and –1.5.

7. Assuming that relative prices do not change, if world income grows by 5 per cent per year and the income elasticity of demand for exports of a country is 1.5, its growth of exports will be:
a) 10 percent per year;
b) 7.5 percent per year;
c) 5 percent per year;
d) 2.5 percent per year.

8. The world consists of only two trading groups, the North and the South. The North's income elasticity of demand for exports is 2 and its income elasticity of demand for imports is 1. This means that the South's income elasticity of demand for exports is 1 and its income elasticity of demand for imports is 2. If the North's growth of output is 5 per cent per year, the growth of the South consistent with balance-of-payments equilibrium will be:
a) 7.5 per cent per year;
b) 5 per cent per year;
c) 3.5 per cent per year;
d) 2.5 per cent per year.

9. Suppose that a country's share of its imports in its output is one-third and there is a depreciation of its nominal exchange rate by 6 per cent per annum. If, as a result, workers successfully bargain to avoid any fall in their real wage, this will eventually increase the rate of growth of domestic prices by:

a) 6 percentage points per year;
b) 4 percentage points per year;
c) 2 percentage points per year;
d) 0 percentage points per year.

10. One-quarter of the exports of a country were initially high-tech with an income elasticity of demand for exports of 2. The remaining three-quarters of its exports were basic goods with an income elasticity of 1. With development, the high-tech industries now comprise three-quarters of exports, and the basic industries one-quarter. The income elasticity of demand for imports is 1 in both periods. If the growth of world income is 5 per cent per year, the change in the composition of exports will raise the balance-of-payments growth rate by:

a) 1 percentage point per year;
b) 1.75 percentage points per year;
c) 2.5 percentage points per year;
d) 3 percentage points per year.

17

European monetary union

Sergio Rossi

OVERVIEW

This chapter:

- presents the history of European monetary union, starting with the Werner Plans discussed in 1970 and then focusing on the Delors Plan (of 1989), which led to both the Maastricht Treaty signed in 1992 and the adoption of the euro, a single European currency, in 1999;
- explains the institution and workings of the European Central Bank, including its governance and lack of accountability with regard to its single monetary policy goal;
- focuses on the euro area crisis, to explain its monetary and structural factors, thus showing the fundamental flaws of the European single-currency area as well as of its anti-crisis policies at both national and European levels;
- proposes an alternative path to European monetary integration, in the spirit of John Maynard Keynes's International Clearing Union based on a supranational currency unit, which can be issued without the need for member countries to dispose of their monetary sovereignty and thereby preserves national interest rate policies as a relevant instrument to steer economic performance at the euro area level;
- points out the critiques raised by Robert Triffin against a single-currency area for Europe and discusses his own alternative proposal in that regard.

Readers will thereby understand the structural and institutional origins of the ongoing euro area crisis.

KEYWORDS

- **Euro area crisis:** Erupting near the end of 2009, mainstream economists consider it to be a sovereign debt crisis, while many heterodox economists argue that it is a balance-of-payments crisis.
- **European Payments Union:** From 1950 to 1958 a number of European countries took part in a multilateral payment system based on an international standard that they used to calculate their net (debtor or creditor) position once per month.
- **Fiscal consolidation:** A series of austerity measures designed to cut public spending in crisis-hit countries across the euro area, so as to induce economic growth by an increase in financial flows from capital markets.
- **Maastricht criteria:** A series of criteria enshrined in the Maastricht Treaty, in order for would-be euro area member countries to converge nominally in economic terms, so as to be able to adopt the euro in a sustainable way.

Why are these topics important?

This chapter aims to assess the historical and theoretical origins of European monetary union as well as to grasp the structural and institutional factors of the euro area crisis that erupted near the end of 2009, as soon as the then newly elected Greek government discovered and disclosed its actual public debt and deficit figures (much higher than previously announced). The analysis presented in this chapter is relevant to understanding the ideological framework supporting the economic policies adopted by euro area member countries against the crisis, and why these policies are not working as expected within that framework. The chapter also explains the monetary policy of the European Central Bank, both before and after the euro area crisis erupted, thereby providing further important insights into understanding that a radically different series of economic policies is required in order to solve the crisis for good. On these grounds, the chapter also illustrates the workings of a supranational currency in the spirit of Keynes's 'bancor', which should be used by central banks only, to retrieve national monetary sovereignty and thus contribute via interest rate policy to steer the domestic economy. The portrait of Robert Triffin included in this chapter presents his views on European monetary union.

The mainstream perspective

Monetary union is considered as the achievement of fixing the exchange rates of the relevant national currencies, and replacing them with a single currency. The costs to a country of abandoning the single currency are much higher than those for abandoning a fixed exchange rate regime (Box 17.1).

BOX 17.1

SINGLE VERSUS COMMON CURRENCY

A single currency is the only currency that is used by any residents in a given economic space, usually a country. The euro is the European single currency for those agents residing in the euro area, which includes a variety of countries. A common currency, by contrast, is a currency that is used by a variety of agents, without being the only currency available in a given economic space. The bancor proposed by Keynes, for instance, was designed to be a common currency, in the sense that it was meant to be used by national central banks only, because households, firms, banks and the general government sector were going to use their own national currency for the settlement of their own transactions.

In this perspective there would thus be no essential distinction on economic grounds between a fixed exchange rate regime and a single-currency area, both basically forms of monetary union.

In the mainstream perspective, there are two approaches to monetary union. The first ('monetarist') approach considers that by gradually reducing (close) to zero the fluctuation band within which foreign exchange rates may vary over time, the relevant countries can fix these exchange rates definitively and then replace their own currencies with a single currency. Exchange rate fixity would oblige the domestic economy of any country to converge – at least in nominal terms – with regard to the macroeconomic magnitudes of the 'best-performing' country within the single-currency area. In the contrary case, the diverging economies would suffer a loss of competitiveness, hence a likely reduction of economic growth and an increase in unemployment levels. The second ('economist') approach, by contrast, considers that the reduction (to zero) of exchange rate volatility must occur only after the relevant countries have already converged in nominal terms to the level set by the best-performing economic system among them. Otherwise, the pressure to enlarge or to abandon the fluctuation band for the exchange rates of those countries that are having difficulty in abiding by with the convergence process would exacerbate tensions within the fixed exchange rate system, thereby putting the sustainability of the whole monetary union at stake.

Both approaches were considered during the discussions that led to the first and second Werner Plans, presented in 1970 in order to monetarily integrate the (then six) member countries of the European Communities. At that time, the economist approach (mostly favoured by Germany) was

BOX 17.2

BRETTON WOODS AGREEMENT

The Bretton Woods agreement was negotiated and signed in July 1944 by the representatives of 44 countries gathering in Bretton Woods, New Hampshire in the United States. Its aim was to establish a new international monetary regime, after the abandonment of the international gold standard in the 1920s. Two plans were proposed and discussed at the Bretton Woods conference: namely, the Keynes and the White plans. The former was a proposal put forward by the British economist John Maynard Keynes, who suggested to adopt what he named 'bancor', that is, an international currency issued by an international clearing bank for the settlement of international transactions between national central banks, each other agent continuing to use its own national currency as before. The second was a proposal put forward by the American economist Harry Dexter White, who suggested to set up an International Monetary Fund and to put the US dollar at the centre of the international monetary regime, in which the US dollar had a fixed exchange rate with gold, and all other national currencies had a fixed exchange rate with the US dollar. The Bretton Woods agreement encapsulates the White plan.

considered better than the monetarist approach (proposed by France, Italy and Belgium). In spite of this, the European monetary union project was put aside in the 1970s, as a result of the demise of the fixed exchange rate system decided at the Bretton Woods conference (Box 17.2).

In fact, in 1971, US President Richard Nixon announced that the gold convertibility of the US dollar (US$35 for an ounce of gold) was suspended. This led to the abandonment of the Bretton Woods system in 1973, when the international monetary system became in fact a floating exchange rate regime. It was only towards the end of the 1980s that the European monetary union project was resumed, in order to dispose of the asymmetric workings (favouring Germany and the German mark) of the European Monetary System (EMS) set up in 1979, when the European Monetary Cooperation Fund was charged with issuing the European Currency Unit (ECU) in exchange for a share of gold reserves and US dollar deposits provided by the national central banks of those countries that were members of the European Communities at that time.

In 1988, the President of the European Commission, Jacques Delors, chaired the Delors Committee, whose mandate was 'to study and propose concrete stages leading towards economic and monetary union' (Committee for the Study of Economic and Monetary Union, 1989, p. 1). As the Werner Plans

already pointed out, the implementation of a monetary union according to the Delors Committee (p. 15) was to imply three major points:

- total and irreversible convertibility of all national currencies of member countries;
- complete liberalization of capital movements across member countries' borders as well as full integration of their banking systems;
- elimination of fluctuation margins and the irrevocable locking of exchange rates for national currencies of member countries.

These steps were to be accomplished while setting up a European monetary authority, first in the provisional form of a European Monetary Institute (1994–98), then in the form of a European Central Bank (ECB), which was set up on 1 June 1998, and then became the monetary authority of the whole euro area when this was established on 1 January 1999. The ECB was prohibited from financing government spending across the single-currency area, which means that it is not allowed to buy government bonds on the primary market (where governments issue and sell their bonds to finance their spending beyond tax receipts). This institutional set-up was accompanied by a series of convergence criteria that would-be euro area member countries had to respect in order to join the single-currency area. Among these criteria are the two most stringent rules for would-be euro area countries: namely, the upper limit of 3 per cent for the ratio between public deficits and gross domesic product (GDP), and the upper limit of 60 per cent for the ratio between public debts and GDP (see Dafflon and Rossi, 1999). Other criteria included:

- an inflation rate that does not exceed by more than 1.5 percentage points the unweighted average of the inflation rates of the three countries where these rates are lowest across the whole European Union (EU);
- a nominal long-term rate of interest (usually on ten-year government bonds) that does not exceed by more than 2 percentage points the unweighted average of similar interest rates in the three EU countries where the rate of inflation is lowest;
- exchange rate stability, defined as a two-year period during which the exchange rate of the currency concerned should not have been devalued with respect to the euro, and the issuing country should have been participating in the Exchange Rate Mechanism of the EMS without any 'severe tensions'.

A number of would-be euro area countries met these criteria only owing to a variety of accounting fiddles (see Dafflon and Rossi, 1999). Further, once in the euro area, several countries did not respect the two fiscal criteria (as

regards public debts and deficits with respect to the country's GDP), but were not sanctioned as expected, because the sanction procedure requires a decision by the Ecofin Council, which is composed of the finance ministers of the EU member countries. This procedure implies thereby that those in charge of sanctioning a country's government might decide not to do this, in order for them to avoid their own countries' governments being sanctioned if they do not respect the Maastricht fiscal criteria.

Other issues also negatively affect the economic governance of the euro area. In particular, the statute of the ECB does not allow the latter to purchase government bonds on the primary market. This 'no bailout' clause is meant to avoid monetary policy supporting government spending, as it could lead to some inflationary pressures that the ECB must avoid. The ECB mandate is indeed to provide, first and foremost, price stability, which the ECB interprets as making sure that the inflation rate in the euro area remains 'close to, but below, 2 per cent' on a yearly basis. This has led Eurostat (the statistical office of the EU) to set up a Harmonized Index of Consumer Prices (HICP), which should allow the measurement of rates of inflation across the euro area with the same methodology, because national indices of consumer prices differ to a large extent and are therefore largely incomparable (see Rossi, 2001, Ch. 1).

In fact, as the first ECB president explained at the Monetary Dialogue of 17 February 2003, the ECB monetary policy strategy 'basically implies that, in practice, we are more inclined to act when inflation falls below 1% and we are also inclined to act when inflation threatens to exceed 2% in the medium term' (Duisenberg, 2003, p. 10). There is thus asymmetric behaviour of monetary policymakers at the ECB, because they are reluctant to reduce their policy rates of interest until measured inflation has fallen below 1 per cent per annum, while they are willing to increase their interest rates as soon as expected inflation (a virtual magnitude) is close to or above 2 per cent, even though actual measured inflation might be lower. This is owing to the importance given by 'new consensus macroeconomics' to the control of inflation expectations by firms as well as individuals (see Arestis, 2007). The ECB fears that if inflation expectations are higher than 2 per cent, its own monetary policy decisions become more difficult to implement successfully and/or could induce a much higher 'sacrifice ratio', that is, 'the cumulative increase in the yearly rate of unemployment that is due to the disinflation effort divided by the total decrease in the rate of inflation' (Cukierman, 2002, p. 1).

The mainstream view about central bank independence, as a way to avoid government intervention in the decision-making process of monetary policy

BOX 17.3

CENTRAL BANK INDEPENDENCE

Central bank independence may exist in several forms, all of which are meant to free the central bank from any political pressure from the general government sector. Goal independence means that the central bank may choose its own policy objectives, as regards both these objectives and the targeted values for them. Instrument or operational independence means the central bank is free to choose the tools it wants to use in order for it to achieve its policy goals. Financial independence means that the central bank has enough funds to carry on its monetary policy strategy, which includes purchasing a variety of financial assets to achieve its policy goals. Personal independence means that the central bank does not have to obey the orders of the government, particularly as regards the central bank's purchases of government bonds to finance the public sector's deficit.

– thereby providing an institutional framework that should make sure price stability prevails – has made the ECB the most independent monetary authority in the world (Box 17.3).

Central bank independence, however, does not necessarily mean that the central bank abstains from purchasing government bonds on primary markets, if these purchases can support the general economic objectives of public policies. In fact, Article 2 of the Statute of the ECB states that:

> Without prejudice to the objective of price stability, it shall support the general economic policies in the [European] Union with a view to contributing to the achievement of the objectives of the Union as laid down in Article 3 of the Treaty on European Union. (European Central Bank, 2012)

Notably, Article 3 states that the European Union:

> shall work for the sustainable development of Europe based on balanced economic growth and price stability, a highly competitive social market economy, aiming at full employment and social progress, and a high level of protection and improvement of the quality of the environment . . . It shall promote economic, social and territorial cohesion, and solidarity among Member States. (ibid.)

Both before and after the euro area crisis erupted, the ECB has done close to nothing to contribute to the achievement of these objectives. As a matter of fact, in the first decade of the EMU, the ECB policy stance has been characterized by an anti-growth bias: it has been raising interest rates when inflation

BOX 17.4

SOVEREIGN DEBT CRISIS

A sovereign debt crisis is a crisis originating in an excessive public sector debt, which most of time is financed by foreign countries' residents, such as in the case of Greece during the euro area sovereign debt crisis that burst in 2009. In fact, a sovereign debt is a debt of the country as a whole, which consists of both public and private sectors' agents.

expectations threatened to be above 2 per cent, and has been reluctant to cut these rates when measured inflation was below 2 per cent. This attitude also affected the first two years of the euro area crisis, since the ECB initially increased its policy rates of interest twice (in April and July 2011) and then decided to reduce them too late (starting in November 2011) and too slowly, with regard to both the dramatic development of the euro area crisis and the sharp and prompt cut in interest rates decided by the US Federal Reserve since early 2008.

Further, the ECB has been a major actor in imposing 'fiscal consolidation' (that is, austerity) policies to a variety of euro area countries, with a view to them rebalancing their public accounts and thereby disposing of the alleged 'sovereign debt crisis' affecting the euro area negatively (Box 17.4).

These policies stem from the view that the euro area crisis is the result of fiscal profligacy, which has led a number of countries 'to live beyond their means', so that – also to avoid moral hazard behaviour – these countries' governments should reduce their spending to rebalance their budgets in a growth-promoting way (see Mastromatteo and Rossi, 2015). The adjustment process in the public sector of these countries should also concern their private sector, so that 'unit labour costs' (the mainstream measure to assess a country's competitiveness) are reduced by a strong downward pressure on wages, which would allow prices to be reduced on the market for produced goods and services. The reference model is thus a neomercantilist regime for economic growth – as implemented by Germany since the early 2000s – that focuses on the export sector as the major, if not unique, engine of economic growth.

This mainstream view of both the euro area crisis and the most appropriate policies to dispose of it is flawed, however, on both theoretical and empirical grounds. This crisis is not due to fiscal profligacy and therefore cannot be solved with austerity policies. A radically different perspective is required in

order to both understand and eradicate the factors that instigated the euro area crisis much earlier than 2009. These are addressed in the next section.

The heterodox perspective

The roots of the euro area crisis are monetary and structural rather than behavioural. In other words, it was not the behaviour of economic agents in either the public or the private sector that originated such a systemic crisis. In fact, both the institutional setting and the monetary–structural framework of the European single-currency area are deeply flawed. A number of heterodox economists have been pointing this out in different ways since the euro area was designed in the 1990s (see Rossi and Dafflon, 2012). In particular, the 'original sin' of the euro is to be 'a currency without a State' (Padoa-Schioppa, 2004, p. 35). Indeed, Kenen (1969, pp. 45–6) pointed out long ago that '[f]iscal and monetary policies must go hand in hand', which means that both of them must be decided at the same institutional level, and in a cooperative way between the government and the relevant central bank. This is not what occurs in the euro area. The ECB is in charge of monetary policy and, in light of its broad independence, does not consider coordinating its policy stance with national fiscal policies across the euro area. The latter polices, moreover, are much constrained by the Stability and Growth Pact signed in 1997 and by the 'Six-Pack' – measures for economic and fiscal surveillance adopted in 2011 – which amount to balancing the public sector's budget without any leeway to carry out a countercyclical policy aimed at supporting economic growth and thereby contributing to achieving the general objectives of the European Union (see above).

Further, the implications of financialization – that is, a regime where financial markets and financial institutions become prominent with respect to economic activity and the needs of the population – affected the euro area negatively, contributing to the creation of debt-led booms in Spain and Ireland and supporting export-led growth in Germany to levels that were problematic for financial stability (Rossi, 2013). Owing to the ECB single monetary policy, whose 'one-size-fits-all' stance has been a much neglected factor of financial instability across the euro area, trade as well as financial imbalances in that area have been increasing during the first ten years of European monetary union, in both its 'core' countries (such as Germany) and its 'periphery' (the PIIGS, an acronym standing for Portugal, Ireland, Italy, Greece and Spain; see Box 17.5).

As a result, debtor as well as creditor countries within the euro area had been initially profiting from an expanding credit bubble – notably as regards

BOX 17.5

PIIGS

PIIGS is the acronym of Portugal, Italy, Ireland, Greece and Spain. It refers to the countries that were most hit by the euro area crisis that erupted in 2009. This acronym echoes the word 'pigs', that is, animals that are considered to live beyond their own means, thereby profiting from an outsider; in the case of the PIIGS, net saving countries such as Germany.

private sector agents, namely, firms and households – that burst once the consequences of the United States (US) subprime crisis crossed the Atlantic and negatively affected a number of European financial institutions (particularly French and German banks).

As Rossi (2007a, 2013) argues, financialization and free mobility of financial capital across the euro area have increased economic divergence between the core and the periphery of that area. Financial capital has been moving from core countries such as Germany (whose current account surpluses since the early 2000s have been providing a huge amount of savings that local financial institutions lent happily), to peripheral countries such as Spain and Ireland, in which they inflated a nationwide real estate bubble that burst in late 2008 as a result of the global financial crisis induced by the bankruptcy of Lehman Brothers in the United States. The growing unsustainability of these imbalances across the euro area remained unnoticed as long as residents (mainly financial institutions) in creditor countries were willing to lend to private and public sector agents in deficit countries the amounts these needed to finance their debts. Indeed, as Draghi (2014) pointed out, a monetary union implies a transfer mechanism between rich and poor member countries or regions: a viable monetary union includes a fiscal equalization mechanism in order to mitigate the financial differences between these countries or regions, thereby enhancing social cohesion and avoiding secessions (Box 17.6).

Since the euro area lacks such a mechanism, its existence during the first ten years has been made possible through private credit flows from the core to its periphery but at a non-negligible cost: while fiscal equalization implies unidirectional flows that are not remunerated, private credit flows must be reimbursed and are remunerated with an interest rate decided by the creditor (influenced by credit rating agencies), which for the debtor may be or become unsustainable over the long run.

BOX 17.6

FISCAL EQUALIZATION MECHANISM

A fiscal equalization mechanism exists in those countries, such as the United States, Germany and Switzerland, where the federal government level redistributes part of fiscal revenues from the richest to the poorest local governments, in order to reduce the socioeconomic discrepancies among the latter. It is a factor of national cohesion that would also benefit the euro area, if such a mechanism were to be introduced there through a series of fiscal transfers from its richest to its poorest member countries.

BOX 17.7

TARGET2 SYSTEM

The TARGET2 system is the euro area-wide payment system that was set up as a result of European monetary union in 1999. It is meant to carry out international payments across that area, similarly to what occurs within each country through its domestic payments system.

This financial issue has been further aggravated by a monetary problem that remained unnoticed until the euro area crisis erupted near the end of 2009: euro area payments are not final payments for the countries concerned due to the workings of the TARGET2 system, despite the fact that they are final for the relevant residents (see Rossi, 2013; and Box 17.7).

This is because the TARGET2 system does not involve the ECB as a settlement institution for national central banks thereby involved. Generally speaking, in fact, a central bank acts as the settlement institution for those banks that participate in the domestic payments system, because of the book-entry nature of bank money and the fact that no one can finally pay by issuing one's own acknowledgement of debt, which is only a promise of payment (see Rossi, 2007b). A central bank must therefore issue its own means of payment in order for banks to settle their debts finally. This is not, to date, what the ECB has been doing, since it abstains from 'monetizing' the TARGET2 system, which is therefore a clearing system through which national central banks finally never pay (are never paid) their net debtor (net creditor) position.

As a result, TARGET2 balances increased noticeably after the outbreak of the crisis at the end of 2009, mirroring the persistence of trade imbalances

between core and peripheral countries in the euro area. Before the crisis broke out, the savings formed in surplus countries were lent to deficit countries, which therefore issued a number of private or public securities in order for them to borrow these savings. As a result, the TARGET2 balances were kept to a minimum. Since the crisis erupted in 2009, however, residents in creditor countries have stopped buying these securities, as credit rating agencies downgraded them. This induced a dramatic increase in balances within the TARGET2 system, because the savings formed in surplus countries have no longer been lent to deficit countries in the euro area.

The various interventions elicited by the euro area crisis, since the first bailout plan for Greece in 2010, have been trying to address some of the consequences rather than the causes of that crisis, which is actually a balance-of-payments crisis rather than a sovereign debt crisis (see above). In fact, there is no problem *per se* in a trade deficit, as long as the country concerned finally pays this deficit. The problem occurs when trade deficits are not finally paid, as this gives rise to payment deficits, which are the empirical evidence of a monetary–structural disorder, because surplus countries maintain a financial claim over deficit countries, as testified by their unsettled positions within the TARGET2 system, and whose balances therefore accumulate over time (see Rossi, 2012a).

The solution to this problem needs a monetary–structural reform. An orderly working payments system requires a settlement institution for each transaction between any two parties to be paid finally. The characteristic of such an institution is to issue the means of final payment, which is an operation involving the payer, the payee and a go-between playing the role of settlement institution between the two agents involved (Figure 17.1).

Indeed, a final payment happens when 'a seller of a good, or service, or another asset, receives something of equal value from the purchaser, which leaves the

Figure 17.1 The emission of money as a flow

Table 17.1 Loans and deposits resulting from the opening of a domestic credit line

Bank	
Assets	Liabilities
Loan to the payer +£*x*	Deposit of the payee +£*x*

seller with no further claim on the buyer' (Goodhart, 1975/1989, p. 26). In order for this to occur, a bank must issue the number of money units needed to carry out the payment, as a result of which the payer disposes of a purchasing power (in the form of a bank deposit) that is transferred, via the bank or the banking system as a whole, to the payee. The result of this payment is a stock magnitude and is recorded by the bank as shown in Table 17.1.

Starting from a blank slate, to avoid the temptation of assuming the existence of a bank deposit before the relevant payment occurs, one notices that a bank's credit is needed in order for the payer to finally settle their debt against the payee. Something similar is required at the euro area level for those payments that affect the position of national central banks within the TARGET2 system. As Keynes (1980, p. 168) noticed in his own plan presented at the Bretton Woods conference:

> We need an instrument of international currency . . . used by each nation in its transactions with other nations, operating through whatever national organ, such as a Treasury or a central bank, is most appropriate, private individuals, businesses and banks other than central banks, each continuing to use their own national currency as heretofore.

In this passage, the author provides an excellent synthesis between the monetary–structural need to set up an international settlement institution, issuing a means of final payment between those countries (each represented by its own central bank) participating in this international payment system, and the importance of maintaining national currencies in place, as they allow a country to retain its monetary sovereignty and thereby to steer its own interest rates as an important tool to achieve the country's economic policy goals. As Keynes (1980, p. 234) explained, in reforming the international monetary architecture '[t]here should be the least possible interference with internal national policies, and the plan should not wander from the international terrain'. This implies that:

> It is the policy of an autonomous rate of interest, unimpeded by international preoccupations . . . which is twice blessed in the sense that it helps ourselves

> and our neighbours at the same time. And it is the simultaneous pursuit of these policies by all countries together which is capable of restoring economic health and strength internationally, whether we measure it by the level of domestic employment or by the volume of international trade. (Keynes, 1936, p. 349)

Although the Keynes plan was not adopted at the Bretton Woods conference back in 1944, it inspired the setting up of the European Payments Union (EPU) in 1950, as a result of a proposal for a European Clearing Union put forward by Robert Triffin (see Triffin, 1985). As Triffin (1978, p. 15) put it, the 'EPU agreement was a remarkably clean and simple document, embodying sweeping and precise commitments of a revolutionary nature, which overnight drastically shifted the whole structure of intra-European payments from a bilateral to a multilateral basis'. This multilateral basis was provided by the Bank for International Settlements (BIS), which indeed operated as a settlement agent for EPU member countries, each represented by its own central bank. Once per month, the national central banks sent their bilateral payment orders, which the BIS cleared in order to establish a net multilateral (debtor or creditor) position for each of the participating central banks. This position was to be settled in gold or US dollars, if the debtor central bank had already exhausted the initial credit that each country received when the EPU was set up in 1950. This is the weak point of the EPU architecture: the lack of an international settlement institution, issuing its own means of final payment for the participating central banks. The BIS should have been operating not simply as a settlement agent (recording the bilateral and multilateral balances resulting from the EPU operations), but as a settlement institution, whose means of payment is required in order to settle all debts within the EPU.

Indeed, the monetary–structural flaw of the EPU (which was terminated in 1958) was to be reproduced in the TARGET system set up in 1999 to carry out large-value cross-border payments within the Economic and Monetary Union (EMU). As a matter of fact, these payments 'are processed via the national RTGS [real-time gross settlement] systems and exchanged directly on a bilateral basis between NCBs [national central banks]' (European Central Bank, 2007, p. 34). This means that the ECB does not intervene as a settlement institution. In fact, the ECB is just a participating central bank among others in the TARGET2 system, which means that it does not issue central bank money for the settlement of national central banks' (net) positions within that system (see Rossi, 2012a). Hence, these positions remain on the TARGET2 books as long as the underlying flows are not reversed, which means that a net importing (exporting) country should become a

net exporting (importing) country in order for these positions to be settled eventually in real terms.

Now, as a result of the euro area crisis, the rebalancing mechanism introduced by the 'troika' (made up of the ECB, the European Commission and the International Monetary Fund) puts the onus of adjustment on debtor countries only, preserving the creditor countries, in particular Germany, from contributing to this adjustment via an increase in their imports from debtor countries (mostly in the periphery of the euro area). However, as the latter countries suffer from 'fiscal consolidation' policies that reduce the size of their economies, they will not be able to increase their exports very much and will also reduce their imports from creditor countries, as a result of reduced available income, thereby aggravating the situation across the whole EMU rather than contributing to relaunching the euro area economy over the long run. This is why the monetary–structural architecture of the EMU should be reformed in the spirit of Keynes's plan and considering the shortcomings of the EPU as explained above. An international settlement institution for euro area member countries should therefore be set up, with the task of issuing the means of final payment for national central banks – within the TARGET2 system – with the possibility to abandon the single currency, to reintroduce national currencies particularly in those countries where the single interest rate policy of the ECB causes more harm than good.

Let us explore the workings of an international settlement institution for those national central banks that participate in the TARGET2 system. For the sake of simplicity, the ECB should undertake this task and therefore issue the means of final payment between all central banks involved thereby. Figure 17.2 shows the flow of central bank money that the ECB issues whenever there is a payment order involving two national central banks within the TARGET2 system.

Let us call 'international euro' ($€_i$) the international means of final payment issued by the ECB in this perspective. As Figure 17.2 shows, the number of (x) money units issued thereby are credited/debited to the paying as well as the receiving central bank, through an instantaneous circular flow whose result is a stock magnitude entered in Table 17.2.

The double entry in Table 17.2 testifies that the paid central bank has no further claim on the payer central bank, which is the hallmark of a final payment between them. Yet, it should be noted that the credit obtained by the paying central bank should be repaid as soon as it is credited for a commercial or financial export by the residents in its country. This may be facilitated by

Figure 17.2 The emission of central bank money as a flow

Table 17.2 Loans and deposits resulting from the opening of an international credit line

European Central Bank	
Assets	Liabilities
Loan to the paying central bank + €$_i x$	Deposit of the receiving central bank + €$_i x$

the fact that central bank deposits at the ECB are not remunerated: in this case, the paid central bank, acting on behalf of its country, will be induced to spend this deposit as soon as it is recorded, to purchase any financial assets that provide a return or to pay for some commercial imports. In both cases, this can allow deficit countries to finance their trade deficits with a sale of financial assets and to increase their own commercial exports, thus reducing trade imbalances through a symmetric rebalancing mechanism. This would actually be an expansionary process for the whole euro area, whose economic system would thus benefit from it, instead of suffering from the recessionary consequences that fiscal consolidation measures are exerting on euro area debtor countries at the time of writing. It would be a means of replacing austerity and recession with economic growth and solidarity across the euro area (see Rossi, 2012b).

REFERENCES

Arestis, P. (ed.) (2007), *Is There a New Consensus in Macroeconomics?*, Basingstoke, UK and New York, USA: Palgrave Macmillan.

Committee for the Study of Economic and Monetary Union (1989), *Report on Economic and*

Monetary Union in the European Community [Delors Report], Luxembourg: Office for the Official Publications of the European Community.

Cukierman, A. (2002), 'Does a higher sacrifice ratio mean that central bank independence is excessive?', *Annals of Economics and Finance*, **3** (1), 1–25.

Dafflon, B. and S. Rossi (1999), 'Public accounting fudges towards EMU: a first empirical survey and some public choice considerations', *Public Choice*, **101** (1–2), 59–84.

Draghi, M. (2014), 'Stability and prosperity in monetary union', speech at the University of Helsinki, Finland, 27 November, accessed 9 October 2019 at www.ecb.europa.eu/press/key/date/2014/html/sp141127_1.en.html.

Duisenberg, W.F. (2003), 'Monetary dialogue with Wim Duisenberg, President of the ECB', EU Economic and Monetary Committee, Brussels, 17 February, accessed 9 October 2019 at http://www.europarl.europa.eu/RegData/seance_pleniere/compte_rendu/traduit/2003/02-17/P5_CRE%282003%2902-17_DEF_EN.pdf.

European Central Bank (2007), *Payment and Securities Settlement Systems in the European Union, Volume 1: Euro Area Countries*, Frankfurt: European Central Bank.

European Central Bank (2012), 'On the Statute of the European System of Central Banks and of the European Central Bank', *Official Journal of the European Union*, C326/230, accessed 9 October 2019 at https://www.ecb.europa.eu/ecb/legal/pdf/c_32620121026en_protocol_4.pdf.

Goodhart, C.A.E. (1975/1989), *Money, Information and Uncertainty*, 2nd edition, London and Basingstoke, UK: Macmillan.

Kenen, P.B. (1969), 'The theory of optimum currency areas: an eclectic view', in R.A. Mundell and A.K. Swoboda (eds), *Monetary Problems of the International Economy*, Chicago, IL: University of Chicago Press, pp. 41–60.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Keynes, J.M. (1980), *The Collected Writings of John Maynard Keynes, Volume XXV: Activities 1940–1944. Shaping the Post-War World: The Clearing Union*, ed. D.E. Moggridge, London, UK and New York, USA: Macmillan and Cambridge University Press.

Mastromatteo, G. and S. Rossi (2015), 'The economics of deflation in the euro area: a critique of fiscal austerity', *Review of Keynesian Economics*, **3** (3), 336–50.

Padoa-Schioppa, T. (2004), *The Euro and Its Central Bank: Getting United after the Union*, Cambridge, MA: MIT Press.

Rossi, S. (2001), *Money and Inflation: A New Macroeconomic Analysis*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Rossi, S. (2007a), 'International capital flows within the European Monetary Union: increasing economic divergence between the centre and the periphery', *European Journal of Economics and Economic Policies: Intervention*, **4** (2), 309–29.

Rossi, S. (2007b), *Money and Payments in Theory and Practice*, London, UK and New York, USA: Routledge.

Rossi, S. (2012a), 'The monetary–structural origin of TARGET2 imbalances across Euroland', in C. Gnos and S. Rossi (eds), *Modern Monetary Macroeconomics: A New Paradigm for Economic Policy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 221–38.

Rossi, S. (2012b), 'Replacing recession and austerity with growth and solidarity across Euroland', in M. Méaulle (ed.), *Austerity Is Not the Solution! Contributions to European Economic Policy*, Brussels: Foundation for European Progressive Studies, pp. 33–40.

Rossi, S. (2013), 'Financialization and monetary union in Europe: the monetary–structural causes of the euro-area crisis', *Cambridge Journal of Regions, Economy and Society*, **6** (3), 381–400.

Rossi, S. and B. Dafflon (2012), 'Repairing the original sin of the European Monetary Union', *International Journal of Monetary Economics and Finance*, **5** (2), 102–23.

Triffin, R. (1960), *Gold and the Dollar Crisis: The Future of Convertibility*, New Haven, CT: Yale University Press.

Triffin, R. (1978), 'Gold and the dollar crisis: yesterday and tomorrow', Princeton University Essays in International Finance, No. 132.

Triffin, R. (1985), 'Une tardive autopsie du plan Keynes de 1943: mérites et carences' ['A belated autopsy of the Keynes plan: merits and deficiencies'], in A. Barrère (ed.), *Keynes aujourd'hui: théories et politiques*, Paris: Economica, pp. 513–21.

A PORTRAIT OF ROBERT TRIFFIN (1911–93)

Robert Triffin was born in Flobecq, Belgium, on 5 October 1911, and died in Ostend, Belgium, on 23 February 1993. In 1938 he obtained his doctoral degree from Harvard University, USA, where he remained lecturing from 1939 to 1942. He was then appointed to the US Federal Reserve (1942–46), the International Monetary Fund (1946–48), and to the Organisation for European Economic Co-operation (1948–51). In 1951, he became a Professor of Economics at Yale University, USA, where he remained until 1977, when he returned to Belgium and contributed to the debate on European economic integration until his death.

Triffin is best known for his critique on the Bretton Woods system. In 1960 he notably pointed out a paradox (since then known as the 'Triffin dilemma') within that system: the US dollar is an international reserve asset and the United States must therefore run a permanent trade deficit in its current accounts, but this reduces confidence in the US dollar across the world economy, since its exchange rate is weakened as a result of the permanent trade deficit in the US balance of payments (see Triffin, 1960).

Triffin was convinced that the international monetary system must be reformed along the lines of Keynes's plan, and tried to make this more palatable in order to get it realized (Triffin, 1960, p. 71). He repeatedly pointed out that the essential flaw of the Bretton Woods system consists in using national currencies as international reserves (p. 10). He thus proposed a radical reform of international payments, consisting of issuing an international means of payment to be used by national central banks only. This would allow the substitution of national currencies (mainly US dollars and pounds sterling) with a truly international reserve asset in any countries' official reserves (p. 102).

Triffin tried to implement the architecture of Keynes's plan within the framework of the European Payments Union (1950–58), which failed because of the lack of a proper international means of payment as suggested by Keynes.

? EXAM QUESTIONS

True or false questions

1. The European Monetary Union has been set up considering the economic convergence of both real and nominal variables.
2. The European Central Bank has a dual mandate similar to the US Federal Reserve: it aims at price stability as well as maximum employment.
3. The European Central Bank is the settlement institution for national central banks in the TARGET2 payments system.
4. Euro area member countries have abandoned both their fiscal and their monetary sovereignties.
5. National fiscal policies of euro area member countries must respect some common criteria as regards public debt and deficits.
6. The European Central Bank has introduced negative rates of interest in the aftermath of the euro area crisis.

7. The Keynes plan could improve the working of the euro area on monetary–structural grounds.
8. The euro is a currency without a state.
9. The European Central Bank may purchase government bonds on the primary market.
10. The monetary policy of the European Central Bank is not coordinated with national fiscal policies across the euro area.

Multiple choice questions

1. The euro:
   a) is a common currency for euro area member countries;
   b) is a common currency for all European Union countries;
   c) is the single currency of euro area member countries;
   d) is the single currency of all European Union countries.
2. In Figure 17.1, the bank issues money:
   a) with a positive purchasing power;
   b) debiting the payer and crediting the payee;
   c) only if it has some pre-existent savings;
   d) as an asset–liability for every agent involved.
3. In Table 17.1:
   a) the bank can only grant a loan if it has a pre-existent deposit;
   b) the bank creates the deposit with a positive purchasing power;
   c) the bank opens a credit line to the payer, which gives rise to a deposit;
   d) the bank creates the loan *ex nihilo*.
4. The TARGET2 system:
   a) does not settle payments between the participating national central banks;
   b) guarantees payment finality between the participating national central banks;
   c) issues the euro as a common currency to be used by national central banks only;
   d) involves the European Central Bank as an international settlement institution.
5. In the euro area, each country:
   a) must respect the fiscal criteria regarding its public debt and deficit;
   b) is free to decide the amount of its fiscal deficit with regard to GDP;
   c) may receive some fiscal transfers from the European Union;
   d) can decide to revert to its own national currency.
6. Which of the following is not required to become a euro area member country?
   a) International capital mobility.
   b) International labour mobility.
   c) Adoption of the euro as single currency.
   d) Abandonment of monetary policy sovereignty.
7. The euro area:
   a) does not need to coordinate its monetary policy with national fiscal policies;
   b) experienced a financial crisis because of the US subprime bubble;
   c) has an orderly-working payments system between central banks;
   d) needs to be reformed on monetary–structural grounds.
8. Robert Triffin's proposal gave rise to:
   a) the European Central Bank;
   b) the European Payments Union;
   c) the TARGET2 system;

d) none of the above.

9. According to the Keynes plan:

a) each country would retain its own national currency;
b) each country would abandon its own national currency;
c) there would be a single monetary policy across the International Clearing Union;
d) fiscal and monetary policies would be coordinated across the International Clearing Union.

10. An orderly-working monetary union implies:

a) a common currency for its member countries' national central banks;
b) fiscal and monetary policy coordination;
c) fiscal transfers between its member countries;
d) all of the above.

Part V

Recent trends

18

Financialization

Gerald A. Epstein

OVERVIEW

This chapter:

- discusses the concept of financialization and its relevance to understanding the nature and dynamics of our modern capitalist economy;
- explains that financialization has many definitions, but the most commonly used one is that it refers to the increasing role of financial motives, financial markets, financial actors and financial institutions in the operation of the domestic and international economies;
- points out the indicators of financialization in contemporary capitalism, notably the increased size of the financial sector in many countries relative to the overall size of their economies, the increased share of total national income going to people engaged in financial activities, the increased orientation of managers of industrial and other non-financial businesses toward financial activities rather than production activities, continuing or even increased incidence of bouts of financial instability and financial crises such as the global financial crisis that erupted in 2007;
- argues that mainstream economics, at least until recently, has tended to underplay the role of financialization in the macroeconomy, while a growing number of heterodox economists are exploring the nature and macroeconomic implications of financialization, including its impacts on income distribution, financial instability and long-term productivity growth;
- summarizes the big debate about the age and meaning of financialization: is it hundreds or even thousands of years old? Or is it a relatively new phenomenon, 150 years old or less? Relatedly, does it represent a whole new mode of organization for capitalism, or is it simply one

among a number of key institutions and dynamics, along with globalization, digitalization and neoliberalism, which drive our economies today?

KEYWORDS

- **Financialization:** The increasing role of financial motives, financial markets, financial actors and financial institutions in the operation of domestic and international economies.
- **Neoliberalism:** The resurgence of nineteenth-century ideas associated with laissez-faire economic liberalism, which holds that markets should dominate the economy and the government should play a relatively small role within the economic system.
- **Rentiers:** Individuals, businesses or households that get a large share of their income from the ownership of financial assets.
- **Short-termism:** It occurs when economic actors are willing to sacrifice longer-term value for values in the present or very short-term future.

Why are these topics important?

For anyone who did not know it already, the global financial crisis that erupted in 2007 made it completely obvious that finance has become a powerful force in our economy. Unfortunately, most of the mainstream macroeconomics profession, as well as many regulators and practitioners operating in the financial market itself, were largely in the dark on the eve of the crisis. This is because the thrust of the mainstream analysis, in the post-war period, was designed to demonstrate that financial markets were either irrelevant to macroeconomic outcomes or, by facilitating the efficient allocation of resources, a potent force for efficiency and growth (see Chapter 6). The idea that finance could cause the virtual meltdown of the global economy was foreign to their theories and far from their minds.

For heterodox economists, by contrast, at least since the work of the late Hyman Minsky, this fact had been well known (see again Chapter 6). 'Financialization' is the latest, and probably most widely used term by analysts trying to name and understand this contemporary rise of finance and its powerful role. The term was developed long before the crisis of 2007–08 but, understandably, since the crisis hit it has become more popular and widely used.

What is financialization?

As discussed by British economist Malcolm Sawyer (2013), the term 'financialization' goes back at least to the 1990s and probably was originated by the United States (US) Republican political operative and iconoclastic writer Kevin Phillips, who defined financialization as 'a prolonged split between the divergent real and financial economies' (Sawyer, 2013, pp. 5–6).

Scholars have adopted the term, but have proposed numerous other definitions. Sociologist Greta Krippner gives an excellent discussion of the history of the term and the pros and cons of its various definitions (see Krippner, 2005). As she summarizes the discussion, some writers use the term 'financialization' to mean the ascendancy of 'shareholder value', that is, the dominance of stock values as a mode of corporate governance. Some use it to refer to the growing dominance of capital market financial systems where stocks and bond markets are key, as in the United States, over financial systems where banks are more important, as in continental Europe. Some follow the late nineteenth-century and early twentieth-century Austrian Marxist economist Rudolf Hilferding's lead and use the term 'financialization' to refer to the increasing political and economic power of the rentier class and banks. For some others, financialization represents the explosion of financial trading with myriad of new financial instruments. Finally, for Krippner herself, the term refers to a 'pattern of accumulation (that is, saving and investment) in which profit making occurs increasingly through financial channels rather than through trade and commodity production' (Krippner, 2005, p. 14) (Boxes 18.1 and 18.2).[1]

BOX 18.1

GRETA KRIPPNER

Greta Krippner is a US-based sociologist who wrote an early influential article on financialization that was later expanded into a book entitled *Capitalizing on Crisis: The Political Origins of the Rise of Finance* (Harvard University Press, 2011). In this influential work, Greta Krippner traces the longer-term historical evolution that made the rise of finance possible in the United States. She argued that financialization was part of broader transformation of the US economy and not simply due to financial speculation and deregulation. Krippner argues that public policies that created conditions conducive to financialization allowed the state to avoid a series of economic, social and political dilemmas that confronted policymakers as post-war prosperity stalled beginning in the late 1960s and 1970s. In this regard, the financialization of the economy was not a deliberate outcome sought by policymakers, but rather an inadvertent result of the state's attempts to solve other problems.

BOX 18.2

RUDOLF HILFERDING

Rudolf Hilferding, born in 1877, was an Austrian medical doctor who later on developed a profound knowledge and interest in political economy, having become a chief theorist of the Social Democratic Party of Germany and a Finance Minister during the Weimar Republic. A self-proclaimed apprentice and follower of Karl Marx, Hilferding developed, amongst other important Marxist theories, a study on the latest phase of capital development – the work *Finance Capital* (1920/1981). *Finance Capital* pointed out how the banking sector exerted a strong influence over the industry, ultimately leading to the concentration of power in monopolies and cartels, through which war and imperialism would eventually settle in.

Drawing on the development of German capitalism during the late nineteenth century, Hilferding noted that industrial capitalism, facing increased competition, would be forced to obtain credit from the banking system in order to expand its operations, the use of which would result in companies becoming tied to banks. Free competition would then start slowly decaying, which would prompt a higher necessity of employing credit to finance fixed capital in the industrial sector. With the consequent growth of corporations, not only industrial capital would agglomerate, as a similar process of concentration amongst banks would occur, culminating in the development of cartels and monopolies able to impose commercial conditions. The cartelization effect would produce a further amalgamation of banks and industry, the latter being increasingly dominated by finance. As such, a growing share of industrial capital would dissociate from industry and be reallocated into financing operations; those who previously worked in the industrial sector would become credit and capital managers. On the reverse side, banks too would become industrial capitalists, given their control over a once independent sector; a process coined by Hilferding as 'finance capital', through which the financial sector becomes increasingly concentrated and takes control over the national industry.

What all these definitions have in common is the idea that finance has become more important and that it has become a more powerful driver of the economy now than in the past. In order to cast a wide net and incorporate the key ideas in these diverse definitions, it is useful to define the term quite broadly and generally, as 'the increasing role of financial motives, financial markets, financial actors and financial institutions in the operation of the domestic and international economies' (Epstein, 2005, p. 3). This definition focuses on financialization as a process, and is quite agnostic on the issue of whether it constitutes a new mode of capitalist organization or an entirely new phase of capitalism, or whether it is simply one important trend of capitalism along with several other important ones including globaliza-

tion, digitalization and neoliberalism. Broad definitions such as mine have the advantage of incorporating many features but have the disadvantage, perhaps, of lacking specificity. Other analysts have used variations on the term 'financialization' to refer to more or less the same set of phenomena. For instance, Palley (2014, p. 8) uses the term 'neo-liberal financialization' to emphasize the importance of neoliberalism as part and parcel of the rise of financialization. Some have not referred to financialization but to 'finance-dominated capitalism' (Hein, 2012).

How old is financialization?

Another important debate is on the periodization of financialization. Is it only a recent phenomenon – say, important since the 1980s – or does it go back at least 5000 years, as Sawyer (2013, p. 6) has suggested? If it goes back a long time, does it come in waves, perhaps linked with broader waves of production, commerce and technology, or is it a relatively independent process driven by government policy such as the degree of financial regulation or liberalization (see Orhangazi, 2008a, 2008b)? Arrighi (1994) famously argued that over the course of capitalist history, financialization tends to become a dominant force when the productive economy is in decline, and when the dominant global power (or hegemon) is in retreat. A good example of this can be found in the early twentieth century when the United Kingdom (UK) was losing power relative to Germany and the United States, and the UK economy was stagnating. This period was also characterized by a considerable increase in financial speculation and instability.

This raises the question as to whether the current period of financialization is due to the reduced role of the United States in the world economy and the rise of China and India. The connection may seem distant, but some argue that the rise of both China and India means that US and European non-financial capitalists cannot compete as well as before, and therefore are turning to financial speculation because they cannot compete in productive investment.

Dimensions of financialization

If one takes a broad perspective on financialization, then one can identify many dimensions of it. One is the sheer size and scale of financial markets and can be seen quite clearly in the large growth in the size of the financial sector relative to the rest of the economy over the last several decades. This growth in finance has been a quite general phenomenon in many parts of the world. For the most part, this chapter will focus on data from the United States (see Epstein and Crotty, 2013).

The growth of finance relative to the size of the economy since 1980 has been nothing short of spectacular. A few pieces of data illustrate this point well. Let us start with the profits of financial institutions. In the United States, financial profits as a share of gross domestic product (GDP) were around 10 per cent in the 1950s. By the early 2000s, financial profits constituted about 40 per cent of total profits in the United States, a historical high. After a sharp decline during the global financial crisis that erupted in 2007, financial profits have recovered to above 30 per cent of total profits, well above the average for the post-war period.

Naturally, with profits having grown so significantly, the size of the financial sector is likely to have been growing as well. Financial sector assets relative to GDP were less than 200 per cent from 1950 to 1985. By 2008, they had more than doubled to well over four times the size of the economy. After a short dip following the global financial crisis, financial sector assets had grown to almost 500 per cent of GDP by 2015.

Also, trends in the United Kingdom are similar to those in the United States, as regards both the size and the profitability of the financial sector. Its profitability had been growing substantially in the post-war period until the global financial crisis erupted and has resumed its growth since that time. Indeed, since the crisis, the growth of financial assets in the United Kingdom have outpaced those in the United States, Germany and Japan relative to GDP (Lapavitsas, 2013, pp. 205–11). More generally, the size of the financial sector and financial profits relative to the size of the economy has grown substantially in most European countries over this period.

Another dimension that characterizes financialization in many countries has been an increase in the financial activities and financial orientation of non-financial corporations. De Souza and Epstein (2014) present data on the financial activities of non-financial corporations in six financial centres, namely the United States, the United Kingdom, France, the Netherlands, Germany and Switzerland. They show that in all six countries – with the possible exception of France – non-financial corporations significantly reduced their dependence on external borrowing for capital investment. Indeed, in three of these countries – the United Kingdom, Germany and Switzerland – non-financial corporations became net lenders, rather than net borrowers, indicating an increasing role for financial lending as a profit centre for non-financial corporations in these countries. Lapavitsas (2013) showed similar trends for the United States, the United Kingdom, Germany and Japan.

A key aspect of financialization that analysts consider as being particularly pernicious has been the vast increase in debt levels in many countries and

many sectors (Taylor, 2014). In many countries, this debt is taken on by the financial sector itself, and by the real-estate sector. Debt, or leverage, is an accelerator that enables the financial system to generate a credit bubble, thereby allowing some actors (such as private equity and hedge funds) to extract wealth from companies, and that can quicken the pace of economic activity more generally. It is the accelerator on the way down after the bubble bursts, leading to distress, deflation and bankruptcy. On the way up, debt helps to magnify the rate of return on your investment: if you put in $1 of your own investment and borrow $9, and you make a 10 per cent return on the $10 investment, then the dollar you earn is really a 100 per cent return on the dollar you put into the pot. But if this investment loses 10 per cent ($1), you will have lost 100 per cent of the dollar of your own money you put in. So, debt (or leverage) magnifies your gains and losses.

Playing a key role in the development of financialization is the role of financial innovation; that is, the change in financial products, instruments and techniques (Box 18.3).

To be sure, financial innovation has played a key role in the development of recent financial practices that contributed significantly to the massive growth in financial activities, and that ultimately contributed to the financial crisis (Wolfson and Epstein, 2013). Among these key financial innovations have been securitization and structured financial products such as asset-backed securities (ABSs), collateralized debt obligations (CDOs), the growth of credit derivatives, such as credit default swaps (CDSs), which both facilitated and then became embedded in these structured products themselves, and innovation in wholesale funding of banks in short-term global markets (Boxes 18.4, 18.5 and 18.6).

These financial innovations have implications that are global in scope. For instance, the Bank for International Settlements in Basel, Switzerland, reports that the global use and level of trading in these instruments have grown spectacularly over the last several decades (Bank for International Settlements, 2013). This process of financial innovation has clearly helped to drive financialization, both within countries and globally.

Financialization and non-financial corporations

There are other important dimensions of the increased financial activities related to non-financial corporations. Among the most important are the increased role of financial activities as a determinant of the pay packages of top management of non-financial corporations, including, most importantly,

BOX 18.3

FINANCIAL INNOVATION

Financial innovation refers to the process of creating new financial instruments, technologies, institutions and even markets. Broadly speaking, it can be broken down into three specific categories of innovation: financial system or institutional innovations, process innovations, and product innovations.

Institutional innovation refers to the creation of new breeds of financial firms, such as hedge funds, private equity institutions or banks that specialize solely in credit card issuance. These innovations continuously change the landscape of the financial system, from the altering of business structures to a revised legal and supervisory framework, typically following the establishment of new institutions. Process innovation relates to new business practices in financial services, namely those based on technological advances aimed at increasing efficiency. Typical examples include office automation, client data management software, the possibility of online and telephone banking, or any other innovation that allows for a new way of conducting financial transactions. Lastly, product innovation includes, as portrayed by the concept, the introduction of financial products, such as new credit forms, deposits, insurance, derivatives, securitization and other financially complex products, such as asset-backed securities and collateralized-debt obligations.

Financial innovations overall encompass crucial changes on risk management practices and transfer, as well as in credit and equity creation. The advances in the financial system have allowed banks to raise capital in a less costly way, increasing the overall availability of credit in the economy. Nonetheless, these innovations are also linked to the expansion in both complexity and opaqueness of the financial system, increasing financial instability.

the corporate chief executive officer (CEO). Perhaps most important are stock options and other stock-related pay for non-financial corporate management. Stock options allow corporate executives to buy shares of the stock of the company, often at a great discount. The CEO becomes thereby a shareholder, which gives the executives an incentive to pump up the stock price in the short term, buying the stocks with their options, leaving them unconcerned about what happens to the price of the stock (or the value of the company) in the longer term. In the United States, where this is especially prevalent, CEOs on average receive 72 per cent of their compensation in the form of stock options and other stock-related pay (Lazonick, 2014).

This focus on stock prices leads the managers of non-financial corporations to use their revenue to buy back their company's stocks in order to raise stock prices and increase their own compensation. Lazonick (2014) refers to this pressure as leading to management policies of 'downsize and distribute'.

BOX 18.4

ASSET-BACKED SECURITIES

Asset-backed securities (ABSs) are a type of investment security such as a bond or a note, that have the distinctive feature of being collateralized by a pool of financial assets. In a process called securitization, a pool of fixed-income assets such as loans, credit card debts or other receivers are merged together and posteriorly sold to other financial actors as a unique security. This synthetically created security is backed by the underlying assets – that is, the assets become the collateral in case of default – and the respective buyer becomes entitled to the original cash flows from both principal and interest. The security is priced according to the type of the merged assets, and gives the investor a yield of return based on the respective risk. The process of securitization that creates an ABS can be used to transform any revenue-generating financial asset into such security, allowing the issuer to generate liquidity from otherwise illiquid assets that could not be sold independently. Securitization also creates a new market of tradable securities, giving investors more opportunities to invest in a wider variety of income generating assets.

A typical example of an ABS is a mortgage-backed security (MBS), specifically secured by a collection of mortgages obligations. Mortgages are merged into one pool that is further divided into smaller tranches based on the mortgage's inherent risk of default. The tranches are sold to investors as bonds, entitling the buyer to regular returns of interest and principal payments. MBSs played a particularly toxic role during the financial crisis of 2007–08, feeding subprime lending. Subprime mortgages were being pooled and sold as securities, prompting a bubble in real estate prices. As the bubble burst, households defaulted and the MBSs that derived their value from the underlying mortgage plummeted in value.

BOX 18.5

COLLATERALIZED DEBT OBLIGATIONS

Collateralized debt obligations (CDOs) are a more complex version of an ABS that may include mortgages as well as other type of debts, such as bank loan obligations, bonds, cash-market debt instruments, real-estate investment trust debts, or even other CDOs, underlying assets of ABSs, residential or commercial MBSs. Rather than a 'plain vanilla' model, CDOs can be quite complex to track or even understand what compounds their underlying debt. CDOs are issued by special-purpose vehicles created specifically to securitize the debt obligations and are marketed mainly to institutional, rather than retail or smaller investors.

BOX 18.6

CREDIT DEFAULT SWAPS

Credit default swaps (CDSs) are a type of financial derivative that allows for an investor to 'swap' their credit risk with that of another financial actor. They are designed to hedge a credit exposure, typically that of a fixed income product such as a security, between at least two parties. Bonds and other debt securities are amongst the most common hedged instruments through credit default swaps, as they are usually offered with a long-term maturity, which makes it difficult to assess the credit risk over the entire life of the asset. Working as a type of insurance policy for credit exposures, the investor of a CDS commits to make payments to the entity that sold the swap over a specific credit instrument, until the settled maturity date of the contract. In return, the investor is protected against a default of the underlying asset. If such a 'credit event' occurs, the seller of the CDS must pay the investor the value of the hedged asset, along with any interest that would have been paid between the default and the asset's maturity date. Despite defaults being the majority of the credit triggers of traded CDSs, other credit events may include failure to pay, obligation acceleration, repudiation and moratorium. Should any of the credit events occur, the investor may settle the contract and receive the agreed coverage.

This represents a dramatic shift with respect to the earlier strategy of 'retain and reinvest', by which management would retain profits and reinvest them back into the human and technological capital of the firm. The numbers in the case of the United States are staggering. Using a sample of 248 companies that have been listed on the S&P 500 index since 1981, Lazonick (2014) reports that, in 1981, firms used 2 per cent of net income for stock buybacks; between 1984 and 1993, such purchases averaged 25 per cent of net income; from 1994 to 2003, 37 per cent; while in the 2004–13 period they used a full 47 per cent of net income for stock buybacks. Certain large, well-known corporations used an even higher percentage of their income for buybacks.

This focus on stock prices is often seen as a prime example of 'shareholder value' ideology, a perspective considered by some as the very essence of financialization. Shareholder value ideology, promoted in the mainstream of the economics profession by Michael Jensen, among others, argues that since shareholders own the corporations, the goal of the corporation management should be to maximize the corporate value for shareholders. Since, they argue, shareholders bear all the risk in the corporation, then this maximization is the most efficient corporate outcome. However, Lazonick (2014) shows that other stakeholders, such as workers and taxpayers, in fact bear

as much if not more risk than shareholders. Also, Stout (2012) shows that shareholders do not really own the corporations, nor do they all share the same values as embodied in the Jensen ideal. Hence, maximizing shareholder value does not mean maximizing share prices. But it does often lead to short-term, destructive orientation by the company's management.

This is one of the most discussed examples of the role of modern financial markets in creating more short-termism, as a major component of financialization. By 'short-termism' is meant a short-time horizon by economic leaders in making production, investment and financing decisions. This short-termism might lead to underinvestment in long-gestation but highly productive and profitable (in the long run) investments, underinvestment in labour development, underinvestment in research and development activities, and overinvestment in activities that generate short-run profits but that might generate long-run risks and/or losses (Haldane and Davies, 2011). The same kinds of pressures face portfolio managers for pension funds and other institutional investors, leading to a similar focus on short-term returns, sometimes at the risk of longer-term investments.

Evidence of short-termism includes the reduced holding period of equities in financial markets, survey evidence that managers will cut profitable long-term investments to reach short-term profit goals, and the evidence that investors have higher rates of required returns for longer-term investments than is necessary (see Haldane and Davies, 2011).

To summarize, this strand of the literature suggests that financialization not only affects behaviour in the financial sector itself, but also has profound effects on non-financial corporations as well. Let us now turn our attention to the impact of financialization on the household sector.

Financialization and households

As the global financial crisis that erupted in 2007 clearly showed, the process of financialization has not only drawn financial and non-financial institutions into its orbit, but households as well. After all, the epicentre of the financial crisis in the United States was in the home mortgage market, and to some extent one segment of that market, the subprime mortgage market. Lapavitsas (2013) and others have argued that the process of financial incorporation of households led to the 'financial expropriation' of these households by financial businesses, which was most clearly and obviously expressed by the massive loss in housing wealth experienced by poor people and minorities in the United States as a result of the crisis.

The incorporation of households into the circuits of financialization goes beyond the intensive use of mortgage loans to buy homes; sometimes, as we noticed in 2007, with catastrophic consequences. The use of credit cards and other forms of consumer credit, and the widespread indebtedness of students through student loans, also comprise the web of connections that households have come to have with financial markets. In the United States, for example, students have taken on more than US$1 trillion in student loans. This has happened during a period in which young people's employment prospects are still weak, even more so 12 years after the financial crisis erupted at the global level.

All in all, financialization has numerous dimensions and in some countries has moved way beyond the financial sector itself. Financial returns, financial motives, the widespread use of debt, and short-termism, among other aspects, have become crucial, if not dominant, for financial firms, non-financial firms and households. This growth in finance, which accelerated around 1980 in a number of countries, has taken on significant global dimensions as well. The question naturally arises: what is the impact of financialization on the economy and society?

Impacts of financialization

Much of the macroeconomic literature on financialization concerns, of course, the impact of financialization on crucial macroeconomic outcomes such as economic growth, investment, productivity growth, employment and income distribution. The massive literature on the global financial crisis has made it pretty clear that aspects of financialization, including the huge increase in private debt, the use of securitization and complex financial products, the widespread use of complex over-the-counter (OTC) derivatives, and the pernicious fraud and corruption, all contributed to the financial crisis and therefore, quite obviously, undermined financial and economic stability (Box 18.7).

But the impacts of financialization on other macroeconomic outcomes are less obvious, and less studied. Before discussing particular impacts, it will be helpful to present some broad frameworks that have been proposed to understand the impact of financialization on macroeconomics (see Hein, 2012; Palley, 2014).

Macroeconomic models and financialization

Owing to space constraints, I cannot provide here a thorough overview of the rapidly expanding heterodox literature on the models of financialization. Hence

BOX 18.7

SECURITIZATION

Securitization is a financial engineering process through which an institution, known as the originator, pools various financial assets into one group. The group of assets held by the originator is subsequently sold as a package as if only one asset is being transacted, thus removing the merged assets from its balance sheet. The pooling operation typically includes types of contractual debt such as home or commercial mortgages, car or student loans, credit card obligations or any other non-debt asset which generates receivables. The new package is referred to as the reference portfolio, which is subsequently bought by another financial actor – a special-purpose vehicle (SPV), also known as the issuer – who gains ownership and creates tradable securities (hence the term 'securitization'), each representing a stake in the assets of the portfolio. The tradable securities are typically divided into different tranches according to the assets' characteristics, such as type, maturity, interest rate and amount of remaining principal of the loan, all of which reflecting different risk levels and thus different yields. Tranches that carry the higher risk will be compensated by higher returns to attract profit-seeking investors. SPVs can create any desired number of tranches, despite the most common model of securitization involving only three: the senior, encompassing the highest tranche with an investment grade; the mezzanine, or the middle tranche that has an intermediate rating, lower than the investment; and the equity or junior, the lowest tranche, either not rated or rated with a speculative grade. Final investors then buy the securities that correspond to each tranche. When buying a securitized asset, investors take the position of the original holder of the reference portfolio and become entitled to the underlying debt's repayment of principal and interest cash flows.

The process of securitization promotes liquidity creation, as it transforms illiquid assets such as home mortgages into marketable financial assets. It creates a new market for these special types of securities and expands the base of investors, allowing for retail investors to obtain financial assets otherwise restricted to them. Often, the SPV would use an arranger, namely a bank, to distribute the securities over the counter to a higher number of small investors. The SPV leads an investor to assume a creditor role that they are not typically qualified for, increasing the risk of the operation as the investor is not aware of the extent or type of debt that underlies the security.

Securitization became a common practice prior to the Great Recession, where banks typically sold pools of mortgages and personal loans, removing them from their balance sheet and freeing up capital, further allowing them to underwrite additional loans. SPVs would then collect these reference portfolios and sell them in tranches to investors. With the risk of the loans being transferred through securitization, the lending standards of the banks decreased as the amount of loans surged. Furthermore, securitization drastically increased the amount of leverage available and potential transactions the banks could make without any need to raise real economic variables. The increase in financial transactions related to home mortgages fuelled the real estate bubble, but as defaults on mortgages

spread, the tranches whose underlying assets were these mortgages became worthless, prompting a round of defaults and losses amongst the securities' investors. As such, securitization has been pointed out as one of the key reasons for the 2008 Great Recession.

I very briefly discuss one framework only, that of German economist Eckhard Hein. Hein (2012) utilizes a Kaleckian model in which aggregate demand plays a key role in determining both investment and output and where income distribution between profits, wages, and rentier or 'financial incomes' has a big impact on aggregate demand. Hein, for example, identifies three key channels through which financialization can affect macroeconomic variables and outcomes:

1. the objectives of firms and the restrictions that finance places on firms' behaviour;
2. new opportunities for households' wealth-based and debt-financed consumption, and
3. the distribution of income and wealth between capital and labour on the one hand, and between management and workers on the other hand.

With his colleague Till van Treeck, Hein shows that within a Kaleckian framework expansive effects may arise under certain conditions, in particular when there are strong wealth effects in firms' investment decisions and households' consumption decisions. However, they show that even an expansive finance-led economy may build up major financial imbalances – that is, increasing debt–capital or debt–income ratios – which make such economies prone to financial instability.

Financialization and investment

Engelbert Stockhammer (2004) pioneered the theoretical analysis of the impact of financialized managers' motives on investment. He showed that finance-oriented management might choose to undertake lower investment levels than managers with less financialized orientations. The key reason explaining this difference is that managers with financialized orientations have a more short-term orientation than other managers. Stockhammer (2004) presented macro-level econometric investment equations that have been consistent with this impact in several countries.

Özgür Orhangazi (2008b) uses firm-level data to study the impact of financialization on real capital accumulation (investment) in the United

States. He uses data from a sample of non-financial corporations from 1973 to 2003 and finds a negative relationship between real investment and financialization. Orhangazi explains these results by exploring two channels of influence of financialization on real investment. First, increased financial investment and increased financial profit opportunities may have crowded out real investment by changing the incentives of firm managers and directing funds away from real investment. Second, increased payments to financial markets may have impeded real investment by decreasing available internal funds, shortening the planning horizons of the firm management, and increasing uncertainty.

Davis (2013) provides further evidence of the negative impact of financialization on real investment. She also studies a sample of non-financial firms, showing a significant difference between large and smaller firms in the degree to which they receive financial income as a share of total income. Larger firms appear to be more financialized in this sense. Using a firm-level panel, she investigates econometrically the relationship between financialization and investment, focusing on the implications of changes in financing behaviour, increasingly entrenched shareholder value norms, and rising firm-level demand volatility for investment by non-financial corporations in the United States between 1971 and 2011. Importantly, Davis (2013) finds that shareholder value norms have been associated with lower investment, though this relationship tends to be true primarily of larger firms.

These results are consistent with the concerns expressed by heterodox analysts and others that financialization tends to reduce real investment.

Employment, human capital, research and development, and wages

An increasing chorus of analysts among heterodox economists including Appelbaum and Batt (2014) and Lazonick (2014) as well as the Bank of England economists Haldane and Davies (2011) have expressed concerns that short-termism associated with financialization may be coming at the expense of investments in human capital, research and development, and employment and productivity growth, which could have long-lasting negative impacts on the economy. There is some empirical work that is supportive of these fears. For instance, in a set of surveys of corporate managers, it has been shown that many chief financial officers are willing to sacrifice longer-term investments in research and development and hold on to valued employees in order to meet short-term earnings per share targets. In a panel econometric study using firm-level data, other economists similarly find

that managers are willing to trade off investments and employment for stock repurchases that allow them to meet earnings per share forecasts. Appelbaum and Batt (2014), in a survey of econometric studies of private equity firms, find that especially large firms that use financial engineering to extract value from target companies have a negative impact on investment, employment, and research and development in these companies. In short, there is significant empirical evidence that short-termism and other aspects of financial orientation have negative impacts on workers' well-being, productivity and longer-term economic growth.

Also, as many of these studies emphasize, these activities do not maximize shareholder value, but often increase incomes for some managers and shareholders, partly at the expense of other shareholders of the firms, not to mention their stakeholders, such as workers and taxpayers.

Income distribution

The analysis presented so far raises the issue of the overall impact of financialization on income distribution. A key issue in this area concerns the origin of financial profits (see Pollin, 1996). Are they the result of a provision of services by finance to the rest of the economy, as is asserted by most mainstream economic theory? Or does much of these profits come in this era of financialization from the extraction of income and wealth by finance from workers, taxpayers, debtors and other creditors? Levina (2014) proposes that much of financial income comes from access to capital gains in financialized markets and therefore does not necessarily reflect a zero-sum game, as is implied by those who argue that financial returns are extracted rather than result from increased wealth. This issue of the source of financial income is extremely difficult to sort out theoretically, and there is no consensus on this topic (see Lapavitsas, 2013).

There has been some empirical work to look at the impact of financialization on income and wealth distribution. Descriptive analysis about the United States indicates that the top earners (the 1 per cent, or even the 0.01 per cent, of the income distribution) get the bulk of their incomes from CEO pay or from finance. Econometric work looking at the relationship between financialization and inequality is also growing. Tomaskovic-Devey and Lin (2011) present an econometric model indicating that since the 1970s, between US$5.8 trillion and US$6.6 trillion were transferred to the financial sector from other sectors in the economy, including wage earners and taxpayers.

Lin and Tomaskovic-Devey (2013), using a sectoral econometric analysis for the United States, find that in time-series cross-section data at the indus-

try level, an increasing dependence on financial income, in the long run, is associated with reducing labour's share of income, increasing top executives' share of compensation, and increasing earnings dispersion among workers. They carry out a counterfactual analysis that suggests that financialization could account for more than half of the decline in labour's share of income, 9.6 per cent of the growth in officers' share of compensation, and 10.2 per cent of the growth in earnings dispersion between 1970 and 2008. Dünhaupt (2013) finds a negative relationship between financialization and the labour share of national income in a larger set of countries from 1986 to 2007.

Financialization and economic growth

As the massive recession stemming from the global financial crisis that burst in 2007 makes clear, there is no linear relationship between the size and complexity of financial markets and economic growth. Several econometric studies have suggested an inverted U-shaped relationship between the size of the financial sector and economic growth. A larger financial sector raises the rate of economic growth up to a point, but when the financial sector gets too large relative to the size of the economy, economic growth begins to decline (see, for example, Cecchetti and Kharoubi, 2012). To the extent that this relationship is true, economists are still searching for an explanation. One argument is that as the financial sector increases in size, because of its relatively high pay levels, it pulls talented and highly educated employees away from other sectors that might contribute more to economic growth and productivity. As a university professor teaching economics since the 1980s, I can testify that many of my undergraduate students had the dream of going to work on Wall Street. Perhaps some of them could have contributed more elsewhere.

Conclusion

There is little doubt that the size and reach of financial activities, markets, motives and institutions have grown enormously in the last 30 years, relative to other aspects of the economy. There is a great deal of historical and empirical evidence that, at least to some extent, this growth has contributed to economic instability, an increase in inequality, and perhaps to a decline in productive investment and employment relative to what might have occurred otherwise. There is less consensus on whether this constitutes a new epoch, phase or mode of accumulation or what exactly is causing this shift: is it underlying problems in the productive core of the economy, a reaction to broader shifts in the global economy associated with globalization, technological changes associated with digitization, or primarily due to financial deregulation as being part and parcel of neoliberalism?

To some extent, public policies or programmes aimed at reducing the deleterious consequences of financialization depend on the underlying causes of the negative aspects of financialization. If the problem primarily stems from issues of financial regulation, then adopting strict financial regulations as suggested by many, imposing a financial transactions tax to reduce short-term trading, prohibiting destructive stock buybacks, breaking up large banks, changing corporate governance so that corporations take into account the preferences of stakeholders, and a host of other reforms, could well go a long way to taming financialization.

If the problems stem largely from the vast and growing inequality of income and wealth distribution and the political power that this inequality buys, then deeper reforms of taxation, wages and ownership, as well as money in politics, must be implemented. If the problem goes deeper to the underlying capitalist dynamics that lead to financialization, then one must look at even more fundamental reforms.

NOTE

1 See also Lapavitsas (2013) and Orhangazi (2008a, 2008b) for important contributions in this regard.

REFERENCES

Appelbaum, E. and R. Batt (2014), *Private Equity at Work: When Wall Street Manages Main Street*, New York: Russell Sage Foundation.

Arrighi, G. (1994), *The Long Twentieth Century: Money, Power and the Origins of Our Times*, London, UK and New York, USA: Verso Books.

Bank for International Settlements (2013), *Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity*, Basel: Bank for International Settlements.

Cecchetti, S. and E. Kharoubi (2012), 'Reassessing the impact of finance on growth', Bank for International Settlements Working Paper, No. 381.

Davis, L. (2013), 'Financialization and the non-financial corporation: an investigation of firm-level investment behavior in the US, 1971–2011', University of Massachusetts Department of Economics Working Paper, No. 2013-08.

de Souza, J.P.A. and G. Epstein (2014), 'Sectoral net lending in six financial centers', Political Economy Research Institute Working Paper, No. 346.

Dünhaupt, P. (2013), 'The effect of financialization on labor's share of income', Institute for International Political Economy Berlin Working Paper, No. 17/2013.

Epstein, G.A. (2005), 'Introduction', in G.A. Epstein (ed.), *Financialization and the World Economy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 3–16.

Epstein, G.A. and J. Crotty (2013), 'How big is too big? On the social efficiency of the financial sector in the United States', in J. Wicks-Lim and R. Pollin (eds), *Capitalism on Trial: Explorations in the Tradition of Thomas E. Weisskopf*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 293–310.

Haldane, A.G. and R. Davies (2011), 'The short long', speech at the 29th Société Européenne Universitaire de Recherches Financières Colloquium on 'New Paradigms in Money and

Finance?', accessed 22 October 2015 at www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2011/speech495.pdf.

Hein, E. (2012), *The Macroeconomics of Finance-Dominated Capitalism and its Crisis*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Hilferding, R. (1920/1981), *Finance Capital. A Study of the Latest Phase of Capitalist Development*, London: Routledge & Kegan Paul.

Krippner, G. (2005), 'The financialization of the American economy', *Socio-Economic Review*, **3** (2), 173–208.

Krippner, G. (2011), *Capitalizing on Crisis: The Political Origins of the Rise of Finance*, Cambridge, MA: Harvard University Press.

Lapavitsas, C. (2013), *Profiting Without Producing: How Finance Exploits Us All*, London, UK and New York, USA: Verso Books.

Lazonick, W. (2014), 'Profits without prosperity', *Harvard Business Review*, September, 1–11.

Levina, I.A. (2014), 'A puzzling rise in financial profits and the role of capital gain-like revenue', Political Economy Research Institute Working Paper, No. 347.

Lin, K. and D. Tomaskovic-Devey (2013), 'Financialization and US income inequality: 1970–2008', *American Journal of Sociology*, **118** (5), 1284–329.

Orhangazi, Ö. (2008a), *Financialization and the US Economy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Orhangazi, Ö. (2008b), 'Financialisation and capital accumulation in the non-financial corporate sector', *Cambridge Journal of Economics*, **32** (6), 863–86.

Palley, T.I. (2014), *Financialization: The Economics of Finance Capital Domination*, Basingstoke, UK and New York, USA: Palgrave Macmillan.

Polanyi, K. (1944), *The Great Transformation*, New York: Farrar & Rinehart.

Pollin, R. (1996), 'Contemporary economic stagnation in world historical perspective', *New Left Review*, **219** (1), 109–18.

Sawyer, M.C. (2013), 'What is financialization?', *International Journal of Political Economy*, **42** (4), 5–18.

Stockhammer, E. (2004), 'Financialisation and the slowdown of accumulation', *Cambridge Journal of Economics*, **28** (5), 719–41.

Stout, L. (2012), *The Shareholder Value Myth: How Putting Shareholders First Harms Investors, Corporations and the Public*, San Francisco, CA: Berrett-Koehler Publishers.

Taylor, A. (2014), 'The great leveraging', in V.V. Acharya, T. Beck, D.D. Evanoff, G.G. Kaufman and R. Portes (eds), *The Social Value of the Financial Sector: Too Big to Fail or Just Too Big?*, Hackensack, NJ: World Scientific Publishing, pp. 33–65.

Tomaskovic-Devey, D. and K. Lin (2011), 'Income dynamics, economic rents, and the financialization of the US economy', *American Sociological Review*, **76** (4), 538–59.

Wolfson, M.H. and G.A. Epstein (eds) (2013), *The Handbook of the Political Economy of Financial Crises*, Oxford: Oxford University Press.

A PORTRAIT OF KARL PAUL POLANYI (1886–1964)

Karl Paul Polanyi was a Hungarian-American economic historian, economic anthropologist, political economist, historical sociologist and social philosopher. He is best known for his opposition to traditional economic thought and for his book titled *The Great Transformation* (Polanyi, 1944). Polanyi is also remembered today as the originator of substantivism, a cultural approach to economics that emphasizes the way economies are embedded in society and culture. Polanyi's masterpiece, *The Great Transformation*, became a model for historical sociology, and in recent years has been revisited to explore the impacts and likely trajectory of phenomena such as financial liberalization.

Polanyi was born into a Jewish family in Vienna in 1886, at the time the capital of the Austro-Hungarian Empire. He graduated from Budapest University in 1912 with a doctorate in Law. From 1924 to 1933 he was employed as a senior editor of the prestigious magazine the *Austrian Economist*. It was at this time that he first began criticizing the Austrian School of economists, who he felt created abstract models that lost sight of the organic, interrelated reality of economic processes. After the accession of Hitler to office in January 1933, accompanied by a rising tide of fascism in Austria, he left for London in 1933, where he earned a living as a journalist and tutor and obtained a position as a lecturer for the Workers' Educational Association in 1936. His lecture notes contained the research for what later became his most celebrated work, *The Great Transformation*. However, he would not start writing this work until 1940, when he moved to Vermont in the USA to take up a position at Bennington College, where he completed and published his book to great acclaim in 1944. In the 1950s and 1960s, Polanyi lived in Canada, in Pickering, Ontario, where he died in 1964.

The Great Transformation deals with the social and political upheavals that took place in the United Kingdom during the rise of the market economy and relates these crucially to the rise of fascism in Europe and the causes of the Second World War. Polanyi contends that the modern market economy and the modern nation-state should be understood not as separate elements but as the single human invention that he calls the 'market society'. This creation of the market society is the 'great transformation' that features in his book's title. The great transformation was begun by the powerful modern state, which was needed to push changes in social structure and human nature that allowed for a capitalist economy. For Polanyi, these changes implied the destruction of the basic social order that had reigned throughout all earlier history. Central to the change was that social factors of production such as land and labour would now be sold on the market at market-determined prices instead of allocated according to tradition, redistribution or reciprocity. He referred to this commodification of land and labour as the creation of 'fictitious commodities', because to sell these on the market was not only such a dramatic change from how they were treated for thousands of years, but also because they are so essential to human life. So, in fact, they cannot really be treated as commodities without generating dire consequences. The other

➡

key fictitious commodity of a capitalist economy, according to Polanyi, is money.

This attempt to treat labour, land and money as if they could be entirely governed by the market, rather than embedded in and regulated by the broader social order, results in massive social dislocation and spontaneous moves by society to protect itself. In effect, Polanyi argues that once the free market attempts to separate itself from the fabric of society, social protectionism is society's natural response, which he calls the 'double movement'. Polanyi thus attempted to turn the tables on the orthodox liberal account of the rise of capitalism by arguing that 'laissez-faire was planned', whereas social protectionism was a spontaneous reaction to the social dislocation imposed by an unrestrained free market. The Great Depression was the crisis of the market economy based on these fictitious commodities. And both fascism and the Second World War were partly created by society through the double movement, where society tried to reassert itself and try to protect itself from the ravages of the market. Polanyi argued that ultimately a form of socialism would be a far better way for society to embed and regulate the market forces for social stability and health.

Polanyi's ideas have had a resurgence in response to the rise of neoliberalism and associated free market thinking. Some scholars have argued that Polanyi's analysis helps to explain why the resurgence of free market ideas have resulted in such manifest failures as persistent unemployment, widening inequality, and the severe financial crises that have stressed Western economies over the past 40 years. They suggest that the ideology that free markets can replace governments is utopian and dangerous.

Nobel Prize-winning economist Joseph Stiglitz and Harvard economist Dani Rodrik explicitly mention their intellectual debt to Polanyi. Both scholars have been vocal critics of the new market fundamentalism and emphasize the inherent contradictions of the efforts to create a self-regulating economy. Apart from contributing to academic discussions, Polanyi's ideas can also help in essential ways to inform political platforms that can address contemporary social and economic challenges. Polanyi's argument about the double movement suggests that more control over the market economy by society is necessary to sustain healthy social outcomes. Pursuing the idea that certain things do not belong in the marketplace, such as healthcare and education, and a retreat from extreme forms of subordination to the markets in general, can be important steps in this direction.

EXAM QUESTIONS

True or false questions

1. The term 'financialization' was originally coined by Hyman Minsky.
2. According to the evidence presented by firm-level studies, shareholder value norms did not have any effect on investment decisions of non-financial corporations.
3. Securitization refers to the rebundling of loans with the intent of selling them off to investors through setting up independent special-purpose entities.
4. Financialization is a phenomenon that emerged in the early 2000s.

5. Rising student loans can be considered as one of the dimensions of the financialization process.
6. According to empirical evidence, financialization had a positive impact on top executives' share of compensation, and no effect on earnings dispersion among workers.
7. Financial innovations can help to turn illiquid assets such as mortgages into liquid assets such as securities.
8. Financialization increased the popularity of the 'downsize and distribute' model within non-financial corporations.
9. Maximizing shareholder value refers to the same thing as maximizing share prices.
10. Neoliberal ideology holds the view that the government should play a minimal role in the overall functioning of the economy.

Multiple choice questions

1. Which of the following cannot be considered as an indicator of financialization in contemporary capitalism?
   a) Increased political power of the rentier class.
   b) The dominance of stock values as a mode of corporate governance.
   c) 'Retain and invest' strategy by corporations.
   d) Share buybacks.
2. Which of the following cannot be considered as one of the financial innovations that played a key role in the build-up to the Global Financial Crisis?
   a) Asset-backed securities.
   b) Repurchase agreements.
   c) Collateralized debt obligations.
   d) Credit default swaps.
3. Which of the following indicators is not commonly used to inform about the extent of financialization?
   a) Required reserve ratio.
   b) Financial profits as a share of GDP.
   c) Credit-to-GDP ratio.
   d) Financial sector assets relative to GDP.
4. After the Global Financial Crisis of 2007, financial profits as a percentage of total profits:
   a) recovered to over 50 per cent of total profits;
   b) remain unchanged;
   c) decreased substantially;
   d) none of the above.
5. Which of the following is typically associated with financialization of non-financial corporations?
   a) Non-financial corporations becoming net borrowers.
   b) Share buybacks.
   c) 'Retain and invest' strategy.
   d) Investment in capital stock.
6. Which of the following cannot be considered as an evidence of 'short-termism'?
   a) Investment in a new plant.
   b) Reduced holding period of equities in financial markets.
   c) Share buybacks.
   d) 'Downsize and distribute' governance model.
7. Which of the following does not represent the institutional or regulatory changes that took place in the US financial system prior to the Global Financial Crisis?

a) Increase in household indebtedness.
b) Increase in the number of financial innovations.
c) Deregulation of the financial markets.
d) Tightening of lending standards.

8. Which of the following can be considered as possible ways of taming the negative social and economic effects of financialization?
a) Prohibition of share buybacks.
b) Adopting strict financial regulations.
c) Introducing financial transactions tax.
d) All of the above.

9. Which of the following indicators cannot be used to asses financial (in)stability?
a) Debt-to-income ratios.
b) Debt-to-capital ratios.
c) Investment-to-GDP ratios.
d) Asset prices.

10. Which of the following has not been negatively affected by financialization?
a) Top income shares.
b) Long-run productivity.
c) Income inequality.
d) Student indebtedness.

19

Imbalances and crises

Robert Guttmann

OVERVIEW

This chapter:

- points out the prevalence of imbalances and crises;
- explains the mainstream view of equilibrium and exogenous shocks;
- discusses Marx's overproduction versus Keynes's underconsumption;
- presents Kalecki's cyclical dynamics of growth and distribution;
- focuses on business cycles and credit cycles (Hayek versus Minsky);
- explains external imbalances and adjustment options.

KEYWORDS

- **Cyclical growth pattern:** Rather than moving along a steady-state balanced growth path as claimed by mainstream economists, the economy's growth pattern is inherently subject to cyclical fluctuations. In that context, downturns known as recessions can be seen as rebalancing processes by which the market mechanism imposes its discipline.
- **Excess supplies:** The innate tendency of a capitalist economy to end up driving total supply (and the underlying production capacity) beyond total demand, be it because of overproduction conditions (Marx) or underconsumption limits (Keynes).
- **External adjustment:** In a globalized economy, nations are not just facing internal imbalances to cope with, but are also more than ever exposed to shocks from abroad. More generally, they have to keep their economic and financial interactions with the rest of the world more or less in balance, or they will be vulnerable to painful (external) adjustment via

exchange rate fluctuations, economic policy changes or crisis-mediated restoration of balance.

- **Financial instability:** The cyclical growth dynamics of profit-driven market economies is reinforced by a parallel credit cycle in the course of which excessive levels of indebtedness taken on during the euphoric boom phase end up leaving overextended debtors vulnerable to any disruption of income and credit supplies. When such disruption happens, as is inevitable in the course of the cyclical growth dynamics of the economy, financial crises will typically arise and trigger a recessionary adjustment.

Why are these topics important?

Capitalism has a long record of instability. Market imbalances, whereby demand and supply do not match, build up and then explode into crises when generalized enough or enduring for sufficiently long periods. As a matter of fact, it is often through crises that the economic system ultimately corrects its imbalances when no other (market-driven or policy-induced) adjustment mechanism has been allowed to work. In this sense, imbalances are inherent to the system. This reality, one of gradually unfolding imbalances triggering crisis-induced adjustments, is generally outside the purview of mainstream economists, who see markets as self-adjusting. In their world, imbalances are, if at all present, exceptional and passing events. Luckily we have a great body of heterodox analysis presenting a different view, and one has to draw from these alternative thinkers to explain one of the central characteristics of the capitalist system: its propensity towards imbalances and crises.

The mainstream view of a self-adjusting economy

Mainstream economists have always marvelled at the market mechanism organizing the interaction between supply and demand. Emphasis in economics on the market's self-balancing abilities was already evident from the very beginning, when Adam Smith (1759, 1776) coined the expression 'invisible hand' to describe how markets regulated themselves to the benefit of all. A century later, in the Marginalist Revolution (Jevons, 1871; Menger, 1871/1950; Walras, 1873/1954), this self-regulation capacity of a market economy came to be anchored in the notion of equilibrium, given its primary expression in the famous demand-and-supply 'scissors diagram' by Alfred Marshall (1890).[1] His elegant depiction of the market's balancing process assumed that competition among buyers as well as sellers, coupled with the innate desire on both sides to come to terms, would yield a mutually agreed price that clears the market and leaves every participant in it satisfied.

The mathematical models and/or graphical diagrams, which the standard neoclassical approach has used ever since Alfred Marshall, basically assume away all the forces that may prevent such automatic self-balancing. They treat time as logical time ($t-n, \ldots, t; t+1, \ldots, t+n$) able to move in both directions, presume rational actors in pursuit of maximizing their benefits, or describe markets as 'perfectly competitive' and hence fully flexible. From that perspective it is easy to arrive at 'equilibrium' and see any 'disequilibrium' (that is, an imbalance) as a temporary aberration caused by exogenous forces disrupting the self-regulating market mechanism (such as unions interfering with the labour market and government regulation restraining product markets).

In reality, however, time is historic and as such only moves forward. Human beings are not rational thinking-machines, but make decisions with cognitive biases. And markets are often 'imperfectly competitive' inasmuch as some market participants exercise more market power than others. Under these conditions, 'equilibrium' is a far less assured thing. It might then also be true that, rather than being caused by exogenous factors bent on disrupting the market, imbalances are endemic to how the economic system operates in reality, away from its textbook depiction.

If we were to take such an alternative view, we would concede for instance that economic theory ought to explain why economies periodically go through brutal downturns known as recessions. But the mainstream approach, so wedded to the market's power of quasi-instantaneous self-balancing, would rather insist – as yet another expression of its fundamental belief in equilibrium – on a balanced or steady-state growth path.[2]

More recently, neoclassical economists have gone a bit further in accepting the possibility that there may be periods where exogenous shocks – whether positive or negative, and affecting either the supply side or the demand side – may move the economy temporarily off its equilibrium path. They recognize that such periods may be of more than just passing duration, after which the economy tries to find its way back to a steady state of growth. But this approach – the 'real business cycle' theory (Kydland and Prescott, 1982; Long and Plesser, 1983) – makes cyclical fluctuations an exception rather than the rule. And it focuses exclusively on the 'real' economy of physical production, ignoring the possibility of disturbances from the monetary and financial side.

The heterodox alternative

All the while orthodox economists have insisted on a balanced growth path, the economy has instead performed by moving along a distinct up-and-

down pattern. And this discrepancy between theory and reality has in turn always left space for alternative ways of analysing how a capitalist market economy performs over time. The history of economic thought has in that regard been profoundly shaped by the contributions of two great economists and visionaries, one dying the year the other was born (1883): Karl Marx and John Maynard Keynes. Both men put forth coherent alternatives to the orthodoxy's vision of an automatically self-balancing economy moving along a balanced growth path to emphasize instead that system's innate propensity to move in terms of cycles. Marx identified a systemic propensity for overproduction ('excess supplies') as the main culprit for such recurrent up-and-down fluctuations. Keynes, on the other hand, stressed the inability of the system to generate sufficiently high levels of total spending ('aggregate demand') for full employment to be attained, thus pointing instead to capitalism's underconsumption tendency.

Marx's challenge

Like the great classical political economists, notably Adam Smith and David Ricardo, had done before him, Karl Marx too believed that the profit rate had a tendency in the long run to decline. But unlike his predecessors putting the blame on the growing plentifulness of capital depressing its price, Marx (1867/1992, 1894/1959) provided a more complex, two-pronged explanation.

Marx's argument about tendentially declining profitability is rooted in competition forcing producers to adopt increasingly automated and capital-intensive production methods, as a prerequisite for the productivity gains needed to stay competitive and to lower their unit costs. Once this position is attained, firms can then either lower the sales price of their products (and still maintain their profit margin) or obtain higher profit margins (at the same sales price). But that drive eventually undermines itself to the extent that such automation of production methods shrinks the source of profit: the unpaid labour time extracted from the workforce ('surplus value') relative to the ever-growing capital base. The result is a declining profit rate, because the numerator in the profit rate ratio (that is, profits) cannot keep up with its denominator (capital).

It should be noted here that Marx also identified countertendencies capable of slowing or even reversing the profit rate decline (see Marx, 1894/1959, Vol. III, Ch. 14). To increase the profit rate, capitalists could revert to any or all of these important countertendencies: (1) increasing the rate of exploitation of workers; (2) decreasing the wage; (3) cheapening the plant and

equipment used in production ('constant capital'); (4) relative overpopulation; and (5) foreign trade. During an economic crisis, all these countertendencies are more likely to be used by entrepreneurs, thereby reaffirming the role played by crises as a clearing-out mechanism with which to restore balance.

Marx then pointed to a second, parallel source of intrinsic imbalance that rendered capitalist economies inherently cyclical: the dual and contradictory nature of wages. On the one hand, the wage is the principal source of income for consumption expenditures by households, and the largest part by far of total demand in the economy. But the wage is also a private cost for businesses, often their largest production cost item. Each capitalist wants the wages of their own workers to be as low as possible, while at the same time wishing the wages of all other workers (serving as potential consumers) to be as high as possible.

Barring the social wage component provided for by the government (such as pension fund and health insurance) and collective bargaining agreements with unions, the wage is mostly determined in the labour market as a private cost and hence kept as low as possible, too low in fact to absorb the rising output capacity of firms pursuing increasingly automated production methods in competition with each other. Marx therefore saw overproduction conditions as inevitable, and the primary source triggering crises. By eliminating the weakest firms and forcing surviving producers to slash production volumes, crises would in turn serve to eliminate such conditions of overproduction.

Keynes's insight

Keynes had a different starting point for his discussion of capitalism's propensity for crisis, which ultimately looked at the economy as a whole and so ended up giving us a new macroeconomic level of analysis. Writing during the Great Depression of the 1930s, Keynes (1936) understood that this phenomenon could not be explained by the then dominant market equilibrium framework of Marshall. How was it possible to have a sustained decline in economic activity with mass unemployment for so many years? What had happened to the Marshallian argument in favour of the markets' self-balancing capacity to get us out of the crisis and restore full employment?

Keynes understood correctly that most markets had ceased to function like perfectly competitive markets. More and more product markets had moved from having many small sellers to a structure dominated by fewer larger-

sized firms, with the market power to set prices. Firms in such imperfectly competitive market structures typically set prices as a function of their unit costs (under normal production conditions) to which they would then add a target profit margin.[3] These producers would respond to any shortfall in demand for their products with cutbacks in output rather than engage in self-defeating price wars. One reason for their resistance to lower prices in the face of excess supplies was rigid labour costs as more and more workers came to be covered by collective bargaining agreements negotiated on their behalf by trade unions.

In other words, institutional changes in the *modus operandi* of the economy – more highly concentrated market structures with fewer and larger producers capable of administering their prices, wages set in advance by multi-year contracts – had rendered prices much less likely to fall. But that growing downward rigidity of prices also implied that the clearing of markets in the face of excess supplies, via lower prices encouraging more buying, had ceased to function properly.

In analysing the Great Depression of the 1930s, Keynes recognized the prevalence of such downwardly rigid price behaviour, in particular also extending to labour markets. But Keynes went a step further to figure out why the economy would not recover even after prices had finally come down quite a bit.[4] And here he pointed to a crucial fallacy of composition, where what is good and therefore seems logical for individual actors is not so for the system as a whole. On the contrary, it may actually be disastrous for the whole system. Specifically, individual firms with some degree of price-setting power in imperfectly competitive markets would try to stave off price wars and instead be inclined to reduce their output and employment levels whenever their inventories built up amidst slowing sales volumes. While this response makes sense for any firm trying to preserve profits by cutting costs, it has a negative impact on other actors tied to the firm. Both the workers laid off by the firm as well as its suppliers, who now are selling less to the firm, will lose income. And such income losses will cause these actors to cut back their spending, affecting yet another group of sellers adversely. The negative multiplier effect of spending cutbacks translating into income losses, which trigger further cuts in spending, can overpower the economy as a whole and paralyse it even in the face of falling prices, as happened during the 1930s.

The Keynesian Revolution

The Keynesian Revolution in the aftermath of the Great Depression in the 1930s provided a macroeconomic dimension to economic analysis, rooted

in the realization that what makes sense for individuals alone may prove disastrous for all of them together, so that the economy must be understood as more than just the amalgam of markets grouping together individual buyers and sellers. Keynes took this insight to argue in favour of an extra-market agent, who is bound neither by the profit motive nor by the typical budget constraints, stepping in to pull the economy out of its doldrums by injecting doses of spending. This agent was the government, capable of deficit spending whenever needed to boost total demand. His policy prescription of aiming for greater budget deficits to boost the economy found immediate followers among politicians desperate enough to try something new; as in Roosevelt's New Deal in the United States.

Marx and Keynes pointed to the same contradiction embodied in the dual nature of the wage as largest source of total demand and dominant item of production costs. Keynes argued consistently that the private (profit-motivated) sector was incapable of generating adequate levels of total demand to assure high levels of employment. While such spending additions as more business investment or greater net exports could compensate, the sheer size of household consumption (as a percentage of total spending), and its inherent ceiling by restrictions on wage income, tend to make total demand inadequate (relative to full-employment output) unless boosted by the public sector running larger budget deficits.

To summarize, both Marx and Keynes pointed to the market economy's inability to sustain a balanced growth path on its own. While Marx stressed the system's propensity for overproduction and Keynes highlighted instead its underconsumption tendency, both disputed the mainstream's emphasis on the market's self-balancing powers to maintain continuous equilibrium.

Growth and distribution

If the dual nature of the wage as largest source of aggregate demand and private business costs lies behind the inherent supply–demand imbalance identified by both Marx and Keynes, then it stands to reason that there should be a link between income distribution and the growth dynamics of capitalist economies. I am talking here in particular about functional income distribution, notably the wage share relative to other slices of the income pie going to capital (the profit share and the interest share).[5]

Polish economist Michał Kalecki (1937, 1938, 1942, 1971) tried to reconcile this link by tracing the dynamic relationship between wage shares and profit shares over the growth cycle while also seeking to identify their prin-

cipal determinants. For Kalecki the interaction between these two strategic income shares played a major role in the (ultimately cyclical) growth dynamics. This Kaleckian link between economic growth and income distribution integrates micro-level actions, notably an investment function and mark-up pricing by firms operating in imperfectly competitive market structures, with the macro-level determinants of wage and profit shares to yield a more profound insight into capitalism's inherently complex balancing act between total demand and total supply.[6]

Wage rates and productivity growth

Kalecki's attempt to link economic growth and income distribution points to a crucial relationship highlighted by all three heterodox masters: Marx, Keynes and Kalecki. This is the evolution of the wage rate relative to that of productivity. As crystallized in Keynes's important discussion of the 'efficiency wage', the balanced growth of the economy depends on both of these variables growing pretty much at the same rate over time.

If productivity growth outpaces that of wages over a sustained period of time, then sooner or later productivity-driven supply will shoot beyond wage-dependent total demand. This is exactly what happened in the United States during much of the 1920s and also the 2000s, resulting both times in the spread of overproduction conditions setting the stage for a major crisis, first in October 1929 and then again in August 2007.

If, on the other hand, nominal wages grow faster than productivity for several years, then you are more likely to see unit labour costs getting pushed up. This may very well result in (cost-push) inflation. To the extent that firms cannot fully pass on higher unit costs to higher prices, their profits get squeezed and this will hamper their long-term growth potential. We saw such stagflation conditions take root in the United States during the 1970s, when productivity growth suddenly came to a halt for nearly a decade while nominal wages still rose (albeit less than inflation, thus causing a fall in real wages for much of that troubled period). It is better then to have both wages and productivity grow at pretty much the same rate. When that happens, total supply and demand can expand in balanced fashion at that rate, too. And at the same time, wage and profit shares are thereby also kept relatively stable.

Cyclical growth dynamics

Kalecki argued for the strategic significance of income distribution, in particular the importance of profits in their dynamic interaction with investment,

in terms of a cyclical growth pattern. In other words, the feedback effects between profits and investment spending driving business behaviour played out in terms of a cyclical pattern of ups and downs. In such a cycle, economic growth accelerates at first to reach a peak, at which point the economy turns down until that decline bottoms out and recovery resumes; only for that same cyclical pattern to be repeated again.

Business cycles

When looking at how and why the economy grows in such recurrently cyclical fashion, we see, as both Marx and Keynes did, that recoveries tend to turn into euphoric booms during the course of which demand cannot keep up with supply and the system eventually overshoots. In this context the crisis may then be viewed as a process whereby imbalances, which may have accumulated over time, are forcibly counteracted by actors (producers and consumers) making the necessary adjustments in the face of untenable situations. Businesses may react to a build-up of unsold inventories by cutting back production, even at times prices, to rid themselves of those excess stocks. Consumers may at that point also cut back and so further depress total demand. But the key to this crisis-driven adjustment process is for supply to fall faster than demand.

Credit

While one can argue, with Marx or Keynes, that capitalist economies have an inherently cyclical growth pattern, empirical evidence suggests that these cycles are reinforced by the typically procyclical behaviour of credit. When looking at the historic record of any advanced capitalist economy, one can see that most economic downturns were preceded by sudden outbursts of financial instability negatively affecting funding conditions in the credit system, and so forcing cutbacks in the face of squeezed credit. This is especially true for depressions, a deeper and longer downturn than the regularly recurrent recessions that are part and parcel of the normal business cycle fluctuations. A depression typically starts with a spectacularly violent financial crisis, such as the worldwide banking turmoil of 1873, the crash of October 1929, or the subprime crisis after August 2007.[7]

Inclusion of financial instability in this discussion of business cycles necessitates consideration of credit as a significant macroeconomic force in the process of economic growth. The market economy can achieve higher growth rates (and even temporarily stave off acute overproduction conditions that would have triggered a recession on their own) by making credit easily

available at reasonable interest rates. Such credit allows borrowers to spend beyond their current income level and so relaxes their budget constraint.

Endogenous credit-money

Credit-funded spending used to depend on savings by private agents (households, firms) and their transformation into loanable funds by financial intermediaries (banks) offering surplus-savings units (savers) deposits for their savings and then using these funds to make loans to deficit-spending units (borrowers).

Thanks to Roosevelt's monetary reforms in 1933–35, following the collapse of the gold standard in September 1931, the banking system has been capable of going beyond the simple financial intermediation exercise of mobilizing savings and turning them into loanable funds. These reforms – in particular the Glass–Steagall Act of 1933 and the Bank Act of 1935 – established a more expansive credit system for which these measures provided endogenous creation of an elastic currency (credit-money) within the banking system that has facilitated continuous debt-financing of excess spending.[8]

This fact too has escaped mainstream economists. Their monetary theories (see Friedman, 1956, 1968) presume somehow that we still live in a world of an exogenously fixed money stock under the strict control of the central bank; in other words a vertically sloped money supply curve.

Financial instability

While we have institutionalized endogenous money creation in the banking system tied to credit extension of banks (credit-money) to provide the economy with an elastic currency, this improvement has come at the expense of another major source of imbalances, namely incidences of financial instability. The system of credit-money has made affordable credit much more accessible for the great majority of economic actors. But this improved access to debt has caused a greater danger of financial instability. That kind of situation may arise at any time when debtors find themselves having taken on too much debt relative to their income generation capacity, so that they end up unable to service their debts. Or there may be a sharp decline in the price of financial assets, such as when there is a stock market crash.

But these incidences are often no more than isolated moments of trouble for some; passing events without major repercussions for the overall performance of the economic system as a whole. For them to have a bigger impact

on the overall economy, they have to reach a scale large enough to disrupt credit supplies from the banking system feeding the economy.

Yet we know from observation that almost every cyclical turning point in the economic system has involved a financial crisis pushing the economy into recession. We have to account for this undeniable empirical fact with an appropriate theoretical framework that acknowledges the recurrent pattern of financial instability at or near the peak as a trigger of recessionary adjustment. We have, in the end, two good theories that fit the bill: the Austrian school, and Hyman Minsky's financial instability hypothesis.

The Austrians

The Austrian school (Mises, 1912/1953, 1949; Hayek, 1931) emphasized a credit cycle in the course of which an irresponsibly lax central bank permits a build-up of debt and overspending in the wake of excessively low interest rates. Being spoiled with ample supplies of cheap loans, businesses end up making too many bad investment decisions ('mal-investment') and hence find themselves in a crisis.

When looking at the United States (US) Federal Reserve's recent record of responding to any recession, even a shallow and short one, by pushing for much lower interest rates and maintaining those long into recovery (as in 1982–87, 1991–94 and 2000–04), there is something to be said for that line of argumentation. The Austrians have therefore in recent years found more followers. A more important reason for their recent success, however, is ideological. Theirs is a deeply conservative approach, which emphasizes the danger of letting the government interfere with the market mechanism in a discretionary fashion. They blame credit overextension on mistaken central bank policy rather than viewing it as an inherent part of the business cycle dynamics typifying capitalist economies. While procyclical monetary policy feeding credit overextension at first and then tightening near the peak amidst signs of overheating has been a clear feature of many business cycles in both the United States and elsewhere, the key argument of the Austrians rests confined to a 'policy mistake' that justifies their overall conclusion to keep the role of government (in this case the central bank) in the economy minimal. This ideological preference (and the Austrians are consistently libertarian in their policy conclusions) fails to give sufficient weight to the more profound idea that advanced capitalist economies with highly developed financial systems are inherently subject to credit cycles from which follow a propensity to financial instability.

Minsky's financial instability hypothesis

If we truly want to understand financial crises as a phenomenon endemic to the capitalist system, we must turn to post-Keynesian economists. They are the followers of the 'true' Keynes, in sharp contradistinction to various neoclassical reinterpretations that squeezed this original, at times even subversive, thinker back into the general-equilibrium box.[9] Among post-Keynesians exploring the ties between credit, money and economic activity, Hyman Minsky (1982, 1986, 1992) in particular came to stress the importance of financial instability at the cyclical peak as trigger of downturns. After the events of 2007–08, his focus on financial instability as an endogenous feature of capitalist economies with correspondingly cyclical growth patterns has regained the attention it deserves.

Business cycle dynamics *à la* Kalecki or Keynes are reinforced in both upswing and downswing phases by a parallel credit cycle, whose sharp turning point at the cyclical peak arises when acute explosions of financial instability of sufficient force push the economic system into recessionary adjustment. For Minsky this is likely to happen when a growing number of debtors reach excessive levels of indebtedness during the upswing phase to render them highly vulnerable to any slowdown of income generation.

Minsky's argument rests on distinguishing between three different financing positions, each one comparing current income generation with given levels of debt servicing charges (that is, regular interest payments and timely repayment of principal):

- In the 'hedge finance' position, agents (households, businesses) earn enough income to pay off all of their debts; they face no risk from their indebtedness.
- In the 'speculative finance' position, borrowers have enough income to service their debts. But they no longer can pay these debts off all at once. This is obviously a somewhat riskier position to find oneself in.
- Finally, in the 'Ponzi finance' position, agents have to take on new debt just to service their old debts. The Ponzi position is very dangerous and can get easily out of hand. Once fallen into such a downward spiral of having to borrow more and more, debtors will typically be obliged to take on increasingly short-term debt coming due that much faster. Whereas any disruption of income generation can easily move a debtor from a speculative position to a Ponzi position, even slight declines in income can have a devastating impact on the debt-servicing capacity of Ponzi debtors. Minsky aptly described such a position as one of 'financial fragility'.

Incidences of financial instability can happen any time excessively indebted actors default on their debts and impose losses on their lenders. Such incidences become a financial crisis when there are a lot of overextended Ponzi debtors rendering the entire system fragile. Minsky argued that increased financial fragility of the entire system is built into the cyclical dynamics of capitalist economies.

During upswing phases, when the economy grows rapidly and incomes rise steadily, that favourable position breeds optimism. Widespread belief in a good future drives many actors to borrow more, and their lenders to extend credit quite willingly. As investment bets materialize successfully, borrowers and their creditors are willing to take on a bit more risk in pursuit of still higher returns. Here financial innovations take on an important role to the extent that they make it easier to get more debt and live with higher levels of leverage. As recoveries turn into (debt-fuelled and innovation-driven) booms, widely shared euphoria induces systematic underestimation of risks.

This careless pursuit of quick riches eventually turns many actors into Ponzi debtors, often recognized as such only *ex post*, when hitherto rapid income growth has peaked or even started to decline while previously low interest rates have risen. Near the cyclical peak, rising debt servicing costs clash with sharply decelerating income growth, as falling profit rates and spreading overproduction begin to manifest themselves while demand for credit spikes and/or monetary policy tightens in the face of accelerating inflation rates. This squeeze on overextended debtors yields then an incidence of financial instability that, by demonstrating unmistakably the degree to which the entire system has become fragile, shifts the mood swiftly from euphoria to fear, even panic. In the wake of the subprime crisis, Wall Street has come to call this brutal turning point a 'Minsky moment'. There is considerable disagreement among post-Keynesians as to whether or not the crisis of 2007–08 was a Minsky crisis, not least because it emerged from excessive household debt whereas Minsky himself focused primarily on corporate debt. But the very prominence Wall Street gave to his work, by calling the panic following Lehman Brothers's bankruptcy in September 2008 a 'Minsky moment', showed the need for integrating financial crises into the macroeconomic framework of analysis.

Financial crisis as a rebalancing adjustment mechanism

When a 'Minsky moment' occurs, credit just freezes up. Lack of access to credit prompts desperate debtors to dump their assets into declining markets in a mad scramble for cash to survive. Creditors suffer major losses to the

point where they too may fail. Simultaneous pull-backs by borrowers and lenders trigger recessionary adjustment, which inevitably follows such a credit crunch. In that process, overextended producers slash their output levels and production costs. While these reactions lower total supply, they also depress total demand, especially when they cause lay-offs and lower wage income.

Stabilization requires total supply to fall faster than total demand for excess inventories to be eliminated. Recessionary adjustment also requires de-leveraging across the board, either by writing off bad debts or by getting rid of old debts while abstaining from taking on new debt commitments. Eventually these crisis-induced adjustments will have corrected the underlying imbalances sufficiently to make recovery possible.

Such conditions of re-stabilization will arrive earlier and more surely with the help of what Minsky termed 'Big Bank' (that is, a strong central bank serving as effective lender of last resort) and 'Big Government' (that is, an adequately sized government capable of larger deficit spending). Minsky considered these two institutions of economic policy particularly indispensable in stopping a possibly self-feeding spiral of forced asset sales, debt liquidations, losses and cut-backs, which may easily get out of control and throw the economy into depression: the famous debt-deflation spiral identified by Fisher (1933) as the primary mechanism underlying the Great Depression of the 1930s.

External imbalances and adjustments

So far this discussion has been confined to imbalances and crises on a national scale. Since the early 1980s, however, we have gone through an intense globalization process, which has made a large number of countries more interdependent of each other. What happens in one corner of the world has immediate and direct repercussions for other regions.

International crisis dynamics

We have seen this process also in terms of contagion, where a crisis spreads from the country of origin to neighbours and beyond: the 'LDC debt' crisis of 1982–89 (a debt crisis affecting over 50 less-developed countries); the crisis of the European Monetary System in 1992–93; the Asian crisis of 1997–98 ultimately spreading within a year to Russia, Brazil and Argentina as well; and finally the Great Recession of 2007–09 reaching a truly worldwide dimension after its initial manifestation as a bursting US housing bubble.

In each of these four crises the transmission mechanisms driving contagion from one country to the next were financial in nature, whether in the form of speculative attacks on currencies deemed hopelessly overvalued, capital flight by panic-stricken investors, or worried lenders cutting back their credit facilities. This is not really surprising in light of the fact that finance is today easily the most globalized aspect of economic activity.[10]

Underlying those internationally transmitted financial crises that we have witnessed recurrently since the 1980s are the 'external imbalances'. Those mostly took the form of chronically large trade deficits and/or excessive borrowing from abroad ('capital imports') both of which led to unsustainable balance-of-payments disequilibria that undermined available foreign exchange reserves. In each instance the crisis, whether a currency crisis or a debt crisis or a combination of both, would hit countries suffering excessive balance-of-payments deficits.

The balance of payments

Any national economy's connection to the rest of the world occurs through its balance of payments, which accounts for all the transactions between that country and the others. Trade in goods and services, investment income and unilateral transfers (like remittances, which are funds that immigrant workers send back to relatives in their home countries), are all grouped together in the current account. Capital transfers, including those of financial capital (such as portfolio investments comprising loans or trades in securities), all show up in the capital account. Each of these transactions can be a credit item, as when a country sells (goods, services or assets) to another country and gets paid by it; or a debit item, which arises when a country buys from another country and therefore has to pay the seller abroad. An external deficit (vis-à-vis other countries) arises in whatever transaction category when debit items exceed credit items there, amounting to a net outflow of funds.

The balance of payments should always be balanced. Hence, if a country runs a current account deficit, it will have to finance it by running a capital account surplus, meaning that it will have to be a net seller of assets to investors abroad. This is just another way of saying that, in most instances, it will have to borrow funds (that is, take loans or sell securities) abroad. Whenever its capital account surplus is not large enough to cover its current account deficit, the deficit country has to use its reserves.[11] Crises occur when the debtor nation has accumulated too high a level of external debts owed abroad in the wake of chronic or excessive current account deficits, and/or has depleted its reserves.

Currency depreciation

Normally countries should not let it get to that point, since such crises – the kinds caused by excessively large external imbalances – tend to have an especially dramatic impact on very large numbers of actors within the domestic economy now forced to go through a wrenching readjustment process as rapidly as possible. We therefore have to explore ways in which countries running large and/or chronic current-account deficits can bring them under control and, possibly, even revert back to surpluses for build-up of depleted reserves. There are several possible scenarios for such adjustments, to use the jargon of economists.

The most obvious one is an appropriate change in the exchange rate. Any country running a balance-of-payments deficit will have more money flowing to the rest of the world than coming in from abroad. The outflows represent supply of its currency, whereas inflows constitute demand for its currency. Such excess supplies of its currency in the wake of balance-of-payments deficits will drive down the exchange rate of its currency. As the deficit country's currency depreciates, foreign goods, services and assets all become more expensive, since the same foreign currency amount now requires a bigger amount of the domestic currency to pay for. The opposite is true for the actors abroad, with domestic items having been rendered cheaper for them as they now have to put up only a smaller amount of their strengthened currency for a given domestic price.

In theory this currency depreciation should then eliminate the underlying external deficit, as domestic actors cut back purchases of now costlier products from other countries, while actors abroad are attracted to now cheaper domestic items. In practice, however, this may not work out as smoothly, as these quantity changes in the right direction may be (at least temporarily) outweighed by the cost effects of higher prices on foreign goods, services and assets (Box 19.1).[12]

Changes in policy mix

Another adjustment path may arise for deficit countries through internal changes that rebalance the domestic economy to the benefit of its external balance. Let us start with the Keynesian equation:

$$Y\,(\text{GDP}) = C\,(\text{Consumption}) + I\,(\text{Investment}) + G\,(\text{Government expenditures}) + NX\,(\text{Net exports, meaning Exports } X \text{ minus Imports } M)$$

Then take the other side of the same coin:

BOX 19.1

EXCHANGE RATE REGIMES

We are tracing here an adjustment process rooted in flexible (that is to say, market-determined) exchange rates. While the world's two leading currencies, the US dollar and the euro, float freely in response to market forces, many countries still prefer to manage their exchange rates actively instead. Such exchange rate management can range from hard pegs tying a currency to an underlying anchor (often the US dollar) at a fixed rate to softer arrangements known as 'managed float' whereby a central bank tries to keep currency price fluctuations within an acceptable range. When a central bank opts to manage the exchange rate of the currency under its jurisdiction, it does so by intervening actively in the foreign exchange market using its foreign exchange reserves. A central bank of a chronic deficit country often finds itself in a difficult position as it runs down its reserves, having to buy up the excess supplies of its currency, which arise when the country concerned runs a balance-of-payments deficit. It will face the prospect of rising interest rates from its interventions, a prospect of credit tightening bound to intensify if and when speculative pressure against its overvalued currency mounts. It has historically proven difficult for governments to reset their currency's exchange rates for currencies subject to pegs before concerted speculative attacks forced their hand. The experience of nearly half a century, since the beginning of floating for the US dollar in March 1973, has clearly shown that fixed exchange rates, especially hard pegs, are more prone to currency crises in the form of speculative attacks breaking pegs; while flexible exchange rates cause a different, more gradual type of financial instability from large exchange rate fluctuations prone to overshoot.

$$Y\,(\text{National income}) = C + S\,(\text{Savings}) + T\,(\text{Tax revenues})$$

Since these two equations express Y in different ways, one measured in terms of output produced and the other measured in terms of the income generated by such production, we can put the two together. After some rearranging, we then obtain:

$$X - M = NX = (S - I) + (T - G)$$

From that accounting identity it is easy to discern that the external current account (im)balance NX depends on the positioning of the two internal (im)balances between savings and investments in the private sector and the budget balance in the public sector. Changing these two balances impacts directly on the current account balance.

If you want the current account balance to go down, you can use monetary policy to impact on the private sector balance: higher interest rates, for instance, would probably boost *S* and weigh down *I*. Or you can try to change the public sector balance with a change in fiscal policy, consisting in this case of deficit-reducing efforts to raise more taxes *T* and/or slow government spending *G*. So you can tighten monetary policy or fiscal policy, or pursue a combination of both. But such austerity measures are a hard sell politically, since they slow economic growth, create more unemployment and make income less likely to grow, if at all.

Crisis-induced adjustments

Governments thus often refrain from making these necessary, but politically difficult changes in their policy mix until it is too late, at which point external deficits, left unattended for too long, trigger a crisis of confidence among investors living abroad. Once a country gets thrown into such a crisis, it goes through a fairly radical adjustment induced by precisely that crisis on its own. At that point of retrenchment, firms and consumers scramble to hoard cash for precautionary purposes as they cut back sharply on their spending, making *S* shoot up, while business investment falls sharply (lowering *I*).

Such a sudden private sector reversal to surplus, imposing massive spending reductions amidst recession, may be somewhat mitigated by an opposite movement towards greater deficits in the public sector, much of which is automatic in light of fiscal stabilizers kicking in during a recession to lower *T* (as tax revenues decline automatically with shrinking economic activity) and raise *G* (as some income maintenance programmes shoot up during the downturn, such as unemployment benefits).

To the extent, however, that the private sector adjustment dominates the public sector reaction in the opposite direction, the external sector will find its current account balance (*NX*) improving. This also makes sense in real life, since a downturn at home would lower imports while at the same pushing domestic producers harder to export in compensation for the slowing sales at home. Once again it is through crisis that we get adjustments that should have happened earlier less painfully, but did not (Box 19.2).

Symmetric adjustments

Much of this adjustment among deficit countries, difficult as it is, would be rendered a whole lot easier if surplus countries were to do their bit and adjust

BOX 19.2

THE US FINANCIAL CRISIS OF 2007–09

A good illustration of such crisis-induced sectoral shifts rapidly rebalancing the domestic economy is what happened to the United States in the wake of the 2007–09 crisis, as based on data provided by the US Bureau of Economic Analysis and the Federal Reserve Board. When credit froze amidst a global financial panic in September 2008, worried American households and firms, especially those highly leveraged actors saddled with large debt servicing charges and now suddenly cut off from proper access to credit, scrambled to hoard cash. Their widespread switch to precautionary saving pushed America's personal savings rate from minus 3.1 per cent of gross domestic product (GDP) in the second quarter of 2007 to a positive 6.9 per cent two years later. At the same time, investment was slashed amidst a deepening recession. As a result America's private sector balance ($S - I$) switched dramatically from minus 2.5 per cent in mid-2007 to a positive 7.6 per cent by May 2009. A 10 per cent shift over a couple of years would on its own surely have caused a depression, had it not been for a counteracting shift in the public sector. In addition to automatic stabilizers kicking in from recession-induced declines in tax revenue collections and transfer payment increases (for instance, unemployment benefits), US President Bush launched the US$700 billion Troubled Asset Relief Programme (TARP) in September 2008, which was followed in February 2009 by President Obama's US$787 billion stimulus package. These measures helped to increase the US budget deficit ($T < G$) from an annualized minus 2.4 per cent of GDP in mid-2008 to minus 10 per cent a year later. Considering that those two domestic sectoral shifts did not quite offset each other, America's current account deficit ($M > X$) halved from its pre-recession minus 5.4 per cent of GDP to just minus 2.5 per cent less than a year later.

at the same time in the opposite direction. If a chronic surplus country, such as Germany, lowers its current account surplus by, say, half, there would be that much less of an adjustment to do by the deficit countries in, say, the rest of Europe. Hence, the ideal situation would be to have simultaneous reductions of external imbalances among both sets of countries, the deficit countries as well as the surplus countries (see Chapter 17).

But this is not likely to be the case, since there is a fundamental asymmetry between them. Deficit countries are eventually obliged to bring their external deficits under control, when they have accumulated too much external debt or run out of foreign exchange reserves. Surplus countries, on the other hand, do not have any such structural limitation. They are in a position to accumulate foreign exchange reserves basically forever, as illustrated by China's US$4 trillion in reserves built up since the 1990s. And why would they voluntarily

make themselves weaker, rendering wilfully less effective those comparative and competitive advantages that yielded the chronic external surpluses in the first place? So, in reality, the adjustment burden is usually not distributed even-handedly between deficit countries and surplus countries, placing the burden more on the former than the latter.

The international monetary system

The precise distribution of adjustment burdens depends ultimately on the prevailing international monetary system (IMS) in place to guide monetary flows and financial transactions between countries. The IMS typically comprises the choice of internationally accepted media of exchange, determination of exchange rates, and the modalities of adjustment in the face of external imbalances.

At times we have had IMS arrangements that explicitly provided for a symmetry in adjustment burdens, as was the case during the gold standard period (1879–1931) whose 'specie-flow adjustment mechanism' subjected deficit countries to deflationary shocks in the wake of gold outflows while at the same time imposing inflationary shocks on surplus countries owing to gold inflows. Another example of such symmetry was the European Monetary System (1979–99), which provided explicitly for simultaneous currency devaluations by deficit countries and currency revaluations by surplus countries.

Today's market-regulated multi-currency IMS does not have such symmetry, however. On the contrary, it contains one added bias of asymmetry making life harder for the weaker deficit countries. Still based on the US dollar as the primary form of world money, that system affords the United States an 'exorbitant privilege' as issuer of the world's principal international medium of exchange. Having to supply dollars created inside the US banking system to other countries, the United States has to run external deficits that assure net outflows of dollars to the rest of the world. In that sense it is fair to say that all the other countries automatically finance American balance-of-payments deficits whenever they use these dollars to pay each other or as reserves. This amounts to saying that the United States has the advantage of being able to cover its external deficits by borrowing from the rest of the world in its own currency, indefinitely, at very low interest rates and without strict repayment schedules. Whenever some of its foreign debt comes due, the United States just rolls that debt over by replacing one bundle of paper coming due with another bundle of paper.[13]

In this kind of arrangement, in which the United States is freed from any external constraint and can thus pile up external deficits for quite some time, other non-privileged debtor nations are pushed aside by the world's creditors automatically financing US deficits and so providing less funding support for the other deficit nations. These nations do not therefore have as much space for their deficits to be covered as they would have had without the creditors' preferential treatment of the United States.

Concluding remarks

As should have become obvious from the discussion in this chapter, the capitalist market economy is subject to a variety of built-in imbalances, which play themselves out in typically cyclical fashion. Supply may outpace demand to the point of overproduction (or underconsumption) after which excess capacities have to be cut back. Tenuous links between income distribution and the economic growth process further reinforce the dynamics of cyclical up-and-down fluctuations. Business cycles tend to be reinforced by parallel credit cycles, which feed the overproduction tendency on the upswing and the cutback dynamics following financial crises. As national economies become increasingly integrated amidst accelerating globalization, the international dimension of imbalances and crises gains strategic significance. Here we have to figure out how, amidst growing balance-of-payments disequilibria, needed adjustments often end up only getting triggered in the course of currency crises and crisis-induced changes in the policy mix.

NOTES

1 The Marginalist Revolution of the early 1870s replaced the classical political economy's labour theory of value, an effort motivated not least in response to the critique of capitalism raised by Marx (1867/1992). See Chapter 3.

2 The classic model depicting such a non-cyclical, steady-state growth pattern as capitalism's more or less automatic tendency is the Solow–Swan model, after Robert Solow (1956) and Trevor Swan (1956).

3 This pricing formula is at the heart of heterodox thinking about the modern firm, as so well captured by Michał Kalecki (1942) and Alfred Eichner (1976).

4 In the end, the demand shortfall of the Great Depression was so overwhelming and enduring that prices finally did come down significantly, falling in the United States alone by nearly a quarter from pre-crisis levels over a four-year period.

5 Functional income distribution refers to the relative sizes of the national income shares going to wages, profits or financial capital income (such as interest and dividends), as opposed to personal income distribution measuring what percentage of the total income pie goes to the top 20 per cent as opposed to, say, the lowest 20 per cent, the next-lowest 20 per cent, and so forth.

6 Kalecki's ability to integrate micro-level and macro-level aspects set the stage for later efforts in that direction among post-Keynesian economists, notably Eichner (1976), and more recently the agent-based models (see, for instance, Setterfield and Budd, 2011).

7 A recession occurs when the economy experiences two consecutive quarters of negative GDP growth, meaning that the economy has shrunk during six consecutive months. A depression, on the other hand, would typically last a minimum of two years and push the rate of unemployment above 10 per cent.

8 See Guttmann (1994) for an analysis of the radical implications of Roosevelt's monetary reforms giving rise to a supply-elastic credit-money, as opposed to supply-inelastic commodity-money prevailing under the gold standard. Such elastic currency helped to transform our economy in terms of permitting chronic budget deficits, business funding of large-scale investments in mass-production technology, consumer debt, and coverage of external (balance-of-payments) deficits.

9 Early relevant examples of US post-Keynesian economists are Weintraub (1978), Davidson (1972) and Eichner and Kregel (1975). A good discussion of the mainstream view of money (as exogenous stock) versus the post-Keynesian view of credit-money (as endogenous flow) is provided by Moore (1988).

10 Guttmann (2009, 2016) has analysed the phenomenon of global finance as a crucial new force tying countries more closely to each other but also exposing them together to greater volatility.

11 These reserves consist usually of gold, key currencies (such as US dollars, euros), or Special Drawing Rights (SDRs) issued by the International Monetary Fund (IMF) for official settlements between governments of debtor and creditor nations. If not earned from previous surpluses, reserves can also be borrowed in international money and bond markets or, as a last resort, from the IMF.

12 This delayed impact of currency depreciation on the external balance of a country, thanks to price effects dominating volume adjustments in the face of relatively inelastic trade patterns, is known as the Marshall–Lerner condition, named after Alfred Marshall (1879/1930) and Abba Lerner (1952). Price inelasticities pertaining to foreign trade may be so strong in the short run that the external deficit actually gets worse in the immediate aftermath of depreciation, before gradually improving; the 'J-curve' effect, first identified as such by Magee (1973).

13 For a discussion of the unique 'exorbitant privilege' of the United States thanks to the world-money status of its currency, see Bergsten (1996) and Eichengreen (2007).

REFERENCES

Bergsten, C.F. (1996), *The Dilemmas of the Dollar: Economics and Politics of United States International Monetary Policy*, Armonk, NY: M.E. Sharpe.

Davidson, P. (1972), *Money and the Real World*, New York: John Wiley & Sons.

Eichengreen, B. (2007), *Global Imbalances and the Lessons of Bretton Woods*, Cambridge, MA: MIT Press.

Eichner, A. (1976), *The Megacorp and Oligopoly: Micro Foundations of Macro Dynamics*, Cambridge, UK: Cambridge University Press.

Eichner, A. and J. Kregel (1975), 'An essay on post-Keynesian theory: a new paradigm in economics', *Journal of Economic Literature*, **13** (4), 1293–314.

Fisher, I. (1933), 'The debt-deflation theory of great depressions', *Econometrica*, **1** (4), 337–57.

Friedman, M. (1956), *Studies in the Quantity Theory of Money*, Chicago, IL: University of Chicago Press.

Friedman, M. (1968), 'The role of monetary policy', *American Economic Review*, **58** (1), 1–17.

Guttmann, R. (1994), *How Credit-Money Shapes Our Economy: The United States in a Global System*, Armonk, NY: M.E. Sharpe.

Guttmann, R. (2009), 'Asset bubbles, debt deflation, and global imbalances', *International Journal of Political Economy*, **38** (2), 46–69.

Guttmann, R. (2016), *Finance-Led Capitalism: Shadow Banking, Re-Regulation, and the Future of Global Markets*, New York: Palgrave Macmillan.

Hayek, F.A. (1931), *Prices and Production*, London: Routledge.

Jevons, S. (1871), *The Theory of Political Economy*, London: Macmillan.

Kalecki, M. (1937), 'A theory of the business cycle', *Review of Economic Studies*, **4** (2), 77–97.

Kalecki, M. (1938), 'The determinants of distribution of national income', *Econometrica*, **6** (2), 97–112.

Kalecki, M. (1942), 'A theory of profits', *Economic Journal*, **52** (206–7), 258–67.

Kalecki, M. (1971), *Selected Essays on the Dynamics of the Capitalist Economy, 1933–1970*, Cambridge, UK: Cambridge University Press.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Kydland, F. and E. Prescott (1982), 'Time to build and aggregate fluctuations', *Econometrica*, **50** (6), 1345–70.

Lerner, A. (1952), 'Factor prices and international trade', *Economica*, **19** (73), 1–15.

Long, J. and C. Plesser (1983), 'Real business cycles', *Journal of Political Economy*, **91** (1), 39–69.

Magee, S. (1973), 'Currency contracts, pass-through, and devaluation', *Brookings Papers on Economic Activity*, **1**, 303–25.

Marshall, A. (1879/1930), *The Pure Theory of Foreign Trade, The Pure Theory of Domestic Values*, London: London School of Economics and Political Science.

Marshall, A. (1890), *Principles of Economics*, London: Macmillan.

Marx, K. (1867/1992), *Capital, Volume I*, London: Penguin Classics.

Marx, K. (1894/1959), *Capital, Volume III*, New York: International Publishers.

Menger, C. (1871/1950), *Principles of Economics*, New York: Free Press. First published as *Grundsätze der Volkswirtschaftslehre*, Vienna: Braumüller.

Minsky, H.P. (1982), *Can "It" Happen Again? Essays on Instability and Finance*, Armonk, NY: M.E. Sharpe.

Minsky, H.P. (1986), *Stabilizing an Unstable Economy*, New Haven, CT: Yale University Press.

Minsky, H.P. (1992), 'The financial instability hypothesis', Levy Economics Institute of Bard College Working Paper, No. 74, accessed 22 October 2015 at www.levyinstitute.org/pubs/wp74.pdf.

Mises, L. (1912/1953), *Theorie des Geldes und Umlaufsmittel*, Munich and Leipzig: Duncker & Humblot. Translated as *The Theory of Money and Credit*, New Haven, CT: Yale University Press.

Mises, L. (1949), *Human Action: A Treatise on Economics*, New Haven, CT: Yale University Press.

Moore, B.J. (1988), *Horizontalists and Verticalists: The Macroeconomics of Credit Money*, Cambridge, UK: Cambridge University Press.

Setterfield, M. and A. Budd (2011), 'A Keynes–Kalecki model of cyclical growth with agent-based features', in P. Arestis (ed.), *Microeconomics, Macroeconomics, and Economic Policy*, London and Basingstoke, UK: Palgrave Macmillan, pp. 228–50.

Smith, A. (1759), *A Theory of Moral Sentiments*, Edinburgh: Kinkaid & Bell.

Smith, A. (1776), *The Wealth of Nations*, London: Strahan & Cadell.

Solow, R. (1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70** (1), 65–94.

Swan, T. (1956), 'Economic growth and capital accumulation', *Economic Record*, **32** (2), 334–61.

Tugan-Baranovsky, M.I. (1894), *Industrial Crises in Contemporary England: Their Causes and Influences on the Life of the People*, St Petersburg: Skorokhodov.

Tugan-Baranovsky, M.I. (1898), *The Russian Factory in Past and Present*, St Petersburg: Panteleev.

Tugan-Baranovsky, M.I. (1909), *Principles of Political Economy*, St Petersburg: Slovo.

Tugan-Baranovsky, M.I. (1910), *Modern Socialism in its Historical Development*, London: Swann Sonnenschein.

Tugan-Baranovsky, M.I. (1916), *The Social Foundations of Cooperatives*, Moscow: Kusnerev.

Tugan-Baranovsky, M.I. (1918), *Socialism as a Positive Doctrine*, Moscow: U.R.S.S.

Walras, L. (1873/1954), *Éléments d'Économie Politique Pure*, Paris: Pichon & Durand-Auziat. Translated as *Elements of Pure Political Economy*, Homewood, IL: Irwin.

Weintraub, S. (1978), *Capitalism's Inflation and Unemployment Crisis: Beyond Monetarism and Keynesianism*, Reading, MA: Addison-Wesley.

A PORTRAIT OF MIKHAIL TUGAN-BARANOVSKY (1865–1919)

Mikhail Tugan-Baranovsky was a leading economist and political figure in pre-revolutionary Russia. Of Ukrainian origin, Tugan-Baranovsky early on became a proponent of Legal Marxism, a uniquely Russian interpretation of Marxism, which used Marx's notion of revolutions occurring in the most advanced capitalist nations to argue in favour of capitalist development as a necessary condition for later revolution. In his important contributions *Industrial Crises in Contemporary England* (1894) and *The Russian Factory in Past and Present* (1898), Tugan-Baranovsky argued that industrialization was taking root in pre-revolutionary Russia to the point of even giving rise to the kinds of industrial cycles recurrently found in the far more advanced British capitalist economy.

Tugan-Baranovsky's argument explicitly countered the leading anti-Tsarist movement at the time, the Russian socialists known as *narodniks*, who emphasized the vanguard role of the peasantry. During much of the 1890s the struggle with these agrarian-oriented populists brought Tugan-Baranovsky and the other Legal Marxists in coalition with the revolutionary Marxists grouped around Vladimir Lenin stressing the leadership role of industrial urban working-class elites to justify immediate uprisings. While the cooperation between the Legal Marxists and Lenin's communist vanguards managed to defeat the *narodniks*, their coalition broke apart towards the end of the century when the Legal Marxists sided with a turn among reformist factions of the European Left towards social democracy under Edward Bernstein, while Lenin and his followers tested their revolutionary mettle in Russia's 1905 Revolution.

The political tensions among these three brands of opposition to the ruling order – the agrarian populists known as *narodniks*, the reformists of Legal Marxism, and the revolutionary Communists led by Lenin – shaped Tugan-Baranovsky's entire adult life, culminating in his ill-fated and aborted experience as Secretary of Finance of the Ukrainian People's Republic during the October Revolution of 1917.

More rewarding was his academic career, which left us with a remarkable record of publications in a short period of time. From the very start, Tugan-Baranovsky stood out as an original critic of the Marxist labour theory of value, established himself as one of pre-revolutionary Russia's leading historians of economic thought, and contributed in important fashion to our understanding of the link between distribution and capitalism's cyclical growth dynamics. Later on, when he turned away from Marxism, Tugan-Baranovsky focused on the co-operative movement. He was the teacher of leading Soviet economist Nikolai Kondratiev, who gave us the notion of long waves.

Among Tugan-Baranovsky's most important contributions, apart from his first two books mentioned above, are his extensive critique of neoclassical economic theory as presented in his *Principles of Political Economy* (1909), his highly original, ethically grounded vision of socialism presented *in both Modern Socialism in Its Historic Development* (1910) and *Socialism as a Positive Doctrine* (1918), and his *The Social Foundations of Cooperatives* (1916).

EXAM QUESTIONS

True or false questions

1. Mainstream economists focus principally on capitalism's tendency towards imbalances and crisis-induced adjustments.
2. 'Real business cycle' theory regards cyclical fluctuations as the exception rather than the rule.
3. Karl Marx focused on underconsumption, whereas John Maynard Keynes stressed overproduction.
4. Karl Marx's countertendencies stemming the decline in the profit rate are more likely to arise during boom conditions than crisis conditions.
5. The wage is both a source of effective demand and a private business cost.
6. Michał Kalecki linked economic growth to income distribution.
7. Business cycle fluctuations are reinforced by the credit cycle.
8. Hyman Minsky's financial instability hypothesis stressed an inherent tendency of profit-seeking actors to take on too much debt.
9. In the Keynesian sectoral balances equation, trade deficits and budget deficits go hand in hand.
10. The world-money status of the US dollar imposes more burdens on the United States than advantages.

Multiple choice questions

1. The notion of the 'invisible hand':
   a) was coined by Karl Marx;
   b) refers to the self-balancing capacity of a market-regulated economy;
   c) was at the core of the Marginalist Revolution of the 1870s;
   d) involves governments giving the poor a helping hand through social programmes.
2. Standard neoclassical models have all of the following features except:
   a) logical time;
   b) perfect competition;
   c) downwardly rigid prices;
   d) equilibrium conditions when the quantity demanded equals the quantity supplied.
3. Karl Marx argued all of the following except:
   a) overproduction;
   b) the contradictory dual nature of the wage;
   c) underconsumption;
   d) a tendency for the profit rate to fall.
4. John Maynard Keynes:
   a) focused on explaining the Great Depression of the 1930s;
   b) wanted governments to run balanced budgets;
   c) was more of a microeconomist than a macroeconomist;
   d) emphasized above all the link between income distribution and economic growth dynamics.
5. For the economy to be relatively balanced:
   a) productivity should grow faster than wages;
   b) productivity should grow more slowly than wages;
   c) productivity growth should match wage gains;
   d) the relationship between productivity and wages is not particularly relevant.

6. Depressions have all of the following characteristics except:
a) rising inflation;
b) a longer and deeper downturn than recessions;
c) a major financial crisis at its onset;
d) a debt-deflation spiral.

7. Credit-money:
a) ties money creation to credit extension;
b) makes it more difficult to finance economic activity with debt;
c) turns money into an exogenous stock variable;
d) turns savings into loanable funds for the funding of business investment.

8. Which of the following financing positions accentuates financial fragility?
a) Self-financing of economic activity.
b) Hedge finance.
c) Speculative finance.
d) Ponzi finance.

9. The Austrians (Hayek, Mises, and so on):
a) favoured a strong role for the government in the economy;
b) attributed an innate propensity for financial instability to lax monetary policy;
c) were more worried about high interest rates than low interest rates;
d) blamed financial instability on credit overextension by profit-seeking actors.

10. Which of the following adjustments in the face of external deficits is the easiest?
a) Currency depreciation.
b) Currency appreciation.
c) Interest rate hikes aimed at reducing the private sector deficit ($I > S$).
d) Fiscal policy changes aimed at increasing the budget deficit ($G > T$).

20

Sustainable development

Richard P.F. Holt

OVERVIEW

This chapter:

- presents an overview of why we cannot equate economic growth with development in order to achieve sustainable development;
- provides a new definition of economic development that incorporates not just growth, but also quality of life and sustainability;
- proposes a very different method for development from neoclassical economics, by looking at the roles of all capital stocks for sustainable development, which include not just manufacturing capital but also public, human, environmental and social capitals;
- explains that to achieve sustainable development we must go beyond a weak definition of sustainability to a strong definition where the focus is on a new measurement of the standard of living; a definition that captures a broadly based measurement of social welfare represented by economists such as Amartya Sen.

KEYWORDS

- **Broadly based standard of living:** Overall well-being that goes beyond the assumption that economic growth by itself will bring improvement in quality of life.
- **Capital stocks for sustainable development:** Valuing all capital stocks used in producing income, quality of life and sustainability, which include manufacturing, human, environmental, public and social capitals.
- **Sustainable development:** A broadly based and sustainable increase in the overall standard of living for individuals within a community.
- **Sustainable indicators:** Comprehensive indicators to track changes in the different capital stocks and quality-of-life factors to achieve sustainable development and improve quality of life.

Why are these topics important?

The impact of economic growth on the environment has recently become an important public policy issue. Concerns about global warming, depletion of natural resources and worldwide population size have heightened the public awareness that certain growth patterns might be unsustainable for both society and the environment (Holt, 2005, 2009; Stern, 2007; Holt et al., 2009). However, the issues associated with economic growth are much broader than just whether we have adequate natural resources for long-term growth. The assumption that more economic growth automatically brings a higher quality of life that is sustainable is also being questioned. Keynes (1931/1963) raised this concern in his essay 'Economic possibilities for our grandchildren' by implicitly asking if continuous economic growth does improve our well-being. John Kenneth Galbraith (1958/1969, 1971) posed a similar question by asking whether growth automatically gives us a higher quality of life in an affluent society. Kenneth E. Boulding (1966) in his famous article 'The economics of the coming Spaceship Earth' challenged the goals and values of economic growth and asked: is it desirable? These questions force us to rethink the linkage between economic growth, quality of life and sustainability. They also require us to re-evaluate the relationship between economic growth and economic development, and ask whether they are the same.

Neoclassical economics has traditionally equated economic growth with economic development, with the implicit assumption that benefits of economic growth will 'trickle down' to improve quality of life and provide sustainability (see Solow, 1956; Lucas, 1988; Maddison, 1991; Friedman, 2005). There is also the assumption of a trade-off between equality and economic growth, because economic incentives are needed for economic growth (see Panizza, 2002; Partridge, 2005, 2006).

This chapter questions these assumptions by arguing that our standard of living encompasses much more than just increases in total output and income. For example, if per capita income increases but there is more air pollution, less economic opportunity, and traffic congestion in a community, then the standard of living may have fallen. Sustainable development is different from economic growth in that it covers not just growth, but also quality of life and sustainability (Greenwood and Holt, 2010b). That is, all three are on an equal footing instead of having quality of life and sustainability trickle down from economic growth, as assumed in mainstream economics. Sustainable development means addressing all three jointly and recognizing the inherent linkages and possible trade-offs among them. It also calls for us

to acknowledge the necessity and role of all types of capital for sustainability, instead of just focusing on manufacturing capital as we find in neoclassical economics.

This leads us to a very different way of measuring and defining the standard of living. If the standard of living is defined as a measurement of overall well-being, where economic growth can actually take away from one's standard of living, then we need to detach growth from development and look for a broader definition of well-being.

The structure of this chapter is the following. First, there is an overall discussion of why we need to make a distinction between economic growth and development in pursuing policies for improving quality of life and sustainability. Then the chapter lays the foundation for a new approach to economic development, which incorporates quality of life, sustainability and economic growth. This method is very different from neoclassical economics, which focuses on private manufacturing capital for development. Instead, I will look at the roles of all types of capital stocks for sustainable development and recognize their interdependence and linkages. The different capital stocks include manufacturing, public, human, environmental and social capital. I then give a new definition of the standard of living: a definition that truly captures a broadly based measurement of social welfare represented by economists such as Amartya Sen (1987). Let us start by looking at the neoclassical model for economic development.

The neoclassical model of economic growth and development

Traditionally, neoclassical economics focused on short-run price theory instead of dynamic growth models to explain development. In the 1940s that started to change with the post-Keynesian Harrod–Domar model that incorporated 'growth' through levels of saving and productivity of capital (Harrod, 1939; Domar, 1946). Harrod recognized that growth and cycles in the economy are not separate as many neoclassical economists at that time assumed. By developing a simple single-sector model – driven by a savings function – which depended upon the level of output and investment, in turn driven by the rate of change of output, he was able to formulate an economic growth model where the level of investment could exceed or be below its desired level, thereby causing cyclical fluctuations.

Though not dealing directly with the Harrod–Domar model's concern with economic fluctuations, Robert Solow and Trevor Swan independently devel-

oped an economic growth model that dealt with many of its shortcomings. It was a simple model where economic growth follows a steady-state path. The production function has constant returns to scale, diminishing marginal productivity of capital and labour determined by the growth rates of labour and technology. These are determined exogenously (Solow, 1956; Swan, 1956). A limitation of the Solow–Swan model is having exogenous technological development. In response, endogenous growth theory models were developed where economic growth is the result of endogenous and not external forces (Romer, 1986, 1994; Pack, 1994). Endogenous growth theory allows an increasing variety of quality of machinery and external economies from investment in new capital, which can eliminate diminishing returns (Greenwood and Holt, 2010a, p. 161).

What is evident in all these models, for those who are concerned with sustainable development, is the absence of natural resources as a unique limiting factor. The focus is always on capital, labour and technology. There is the overall view that natural resources and sustainability issues can be dealt with by technological change (Nordhaus and Tobin, 1972), the use of market pricing (Coase, 1960) and allocation of non-renewable resources over time by the right discount rate (Hartwick, 1977, 1989). Solow (1974, p. 11) made it very clear that he did not see limitations of natural resources as an issue for economic growth, when he stated in his 1974 Richard Ely Lecture to the American Economic Association that 'the world, can, in effect, get along without natural resources'. In sum, the neoclassical growth model deals with sustainability by technology, substitution effects and market allocation of natural resources over time with a discount rate. This has led to a debate about what is meant by 'sustainability' between mainstream and heterodox economists.

The debate over 'strong' and 'weak' sustainability

The debate between the 'strong' and 'weak' definitions of sustainability started in 1972 when a Massachusetts Institute of Technology (MIT) research team presented a study titled *The Limits to Growth* (Meadows et al., 1972). By the use of computer modelling, the research team showed different economic growth outcomes with the use of neoclassical economics, which in many cases led to dangerous possible outcomes catastrophic to the Earth's ability to sustain such economic and population growth. A vigorous debate took place between mainstream and heterodox economists over the MIT study, and whether we should use a 'strong' or 'weak' definition of sustainability to avoid the depletion of natural resources to maintain a needed ecological balance (Box 20.1).

BOX 20.1

THE LIMITS TO GROWTH AND THE BEGINNING OF THE DEBATE BETWEEN STRONG AND WEAK SUSTAINABILITY

In 1972, *The Limits to Growth* (Meadows et al., 1972) was published to evaluate the consequences of continued economic growth with the assumptions of neoclassical economics. Many consider this work, along with Kenneth Boulding's (1966) article, 'The economics of the coming Spaceship Earth', as the beginning of a heterodox approach to sustainable development. Using a model they called World3, they added more 'essential' relationships and endogenous variables in understanding the consequence of economic growth on the environment. Their model included not just capital inputs and technology, but also agricultural output, population growth, resource use and environmental pollution. *The Limits to Growth* criticized the mainstream models of Solow (1974, 1986) and Hartwick (1977, 1989) for making important variables such as death and birth rates exogenous, as compared to the World3 model that made them endogenous. For example, by making the population rate growth rate endogenous in the model, this made the mortality rate dependent on access to health care and the environment, which would have an impact on sustainable development. Following the work of Sen, fertility is affected by education, family planning and the overall ecosystem, which in return affects economic output and growth.

The report gave a warning: if things do not change with human behaviour, technology and economic growth patterns, the world economy is heading for a natural disaster. If we continue with the path of exponential growth, then in the future we will have fewer options in adjusting to these natural catastrophes. The report ends with the final prediction of what might happen if we yield to a weak definition of sustainability:

> The last thought we wish to offer is that man must explore himself – his goals and values – as much as the world he seeks to change. The dedication to both tasks must be unending. The crux of the matter is not only whether the human species will survive, but even more, whether it can survive without falling into a state of worthless existence. (Meadows et al., 1972, p. 197)

Most mainstream economists recognize that there are external costs associated with pollution, and that non-renewable resources need to be managed. Solow (1992, 1994), for example, has argued that any depletion of natural resources needs to be matched with increases in manufactured capital. The idea that human-made capital can be substituted for natural capital has led to a heated debate between neoclassical and ecological economists of what is the meaning of sustainable development, which has given us two very different definitions of 'sustainability' (Holt, 2009). The first, championed by neoclassical thinking, has been called the 'weak' definition, where losses of

non-renewable natural capital can be substituted for human-made capital. The second comes from ecological economists. They believe that natural capital needs to be sustained for its unique quality, giving us a 'strong' definition (Holt, 2005, p. 176). A working definition of the 'strong' view comes from the United Nations Brundtland Commission: 'Sustainable development is development which meets the needs of the present without compromising the ability of future generations to meet their own needs' (Brundtland, 1987, p. 7).

This fits nicely with a heterodox approach to sustainable development, which includes eradicating social evils such as unemployment and poverty today, along with saving natural resources for future generations. This is a challenging agenda where present needs to end poverty are eradicated without compromising the well-being of future generations by compromising our natural resources. Ecological economists believe that sustainable development cannot be achieved through endless economic growth, but instead changes with our consumption, production and energy use to protect all types of accumulated capital, particularly social and natural.

The neoclassical approach is very different, where increased economic growth provides needed income for investments in new manufacturing capital and technology that can be used for present and future development. This is represented in the 'weak' definition of sustainability. At the heart of this definition is the assumption that there will be adequate substitutes of human-made capital for natural capital. This has become known as the 'Hartwick–Solow' rule (see Solow, 1974, 1986; Hartwick, 1977, 1989). While they admit that there might be some exceptions to the rule, overall it will hold and allow for sustainable development even if natural resources are depleted. In regard to external costs associated with pollution or global warming, we can turn to taxes (Pigou, 1929) and defined property rights (Coase, 1960) to capture these social costs. Over time, the need for new manufactured capital and the use of taxes and property rights will give incentives to entrepreneurs to develop new capital substitutes and resource-saving technologies. And with a carbon tax or tradable pollution permits we will be able to deal with the external costs associated with environmental problems such as global warming.

Ecological economists have responded by arguing that the environment is not simply an input to the economy. More importantly, the economy exists within society and the environment (see Daly, 1996; Daly and Farley, 2004). A sustainable development approach takes the effects of economic decisions and policies on all capital stocks into consideration, and acknowledges the

unique role that natural capital plays. Unlike other types of capital, there are aspects of certain types of natural capital that cannot be reproduced or protected by private ownership or government actions. The environmental saying 'think globally, but act locally' reflects both the global and the individual mien of natural capital and its importance for sustainability. Recognizing that there are different types of natural capital is important: some are renewable and others are not. Forests and fish are renewable, if well managed. But other forms such as the air we breathe, the water we drink and the waste absorption of the atmosphere and oceans are limited and provide life-giving services that cannot be replaced by any other form of capital.

There is also the question of whether efficiency should be the only or even the primary goal with sustainable development, as we find with the traditional economic growth model. Instead, broader and more encompassing social and environmental goals might be needed (Greenwood and Holt, 2010a). Should efficiency be looked at as a means, or an end in itself, as we find with neoclassical economics when considering sustainable development? Contrary to the neoclassical view, post-Keynesians have argued that efficiency has meaning only in a framework of social goals such as full employment or, in this case, sustainable development (Holt, 2007). They also point out that the neoclassical 'weak' definition of sustainability is void of economic and political power, which can affect how natural resources are used and distributed. If one group can control or limit the use of natural resources and shift environmental and health costs onto another, this can cause unjust harm to poor and powerless groups, besides not allocating resources in the most efficient way (see Galbraith 1958/1969; Holt et al., 2009).

DeGregori (1974, pp. 759–70) points out that with corporate power, as an example, we might accept economic outcomes that we would not in other areas of our lives. He quotes the institutionalist Ayres (1962): 'poisoning one's wife is a mortal sin, whereas poisoning thousands of people by selling adulterated food or drugs is a mere business misadventure'. This follows the concerns of those who hold a 'strong' definition of sustainable development, where the people being 'poisoned' are future generations who have no power to influence the decisions of the present generation (Greenwood and Holt, 2010a, p. 163).

For heterodox economists the lack of realism about the real world and how power impacts upon markets is one of the problems with the 'weak' definition of sustainability. The mainstream view has competitive firms as price-takers, which limit their economic power. But industries associated with energy use

and natural resources are likely to be concentrated, which allow for price discrimination. This means that efforts to use a carbon tax or tradable permits as a way to adjust for social costs associated with pollution becomes difficult, since there are so many structural and institutional factors that can lead to price discrimination in these markets (like financial markets that can speculate on the value of future tradable permits, which makes their prices volatile). Yet, those who advocate the use of market auction permits to deal with global warming assume competition in these markets. The distribution of wealth and power is not central to the neoclassical world view of environmental policies. For heterodox economists it is, as they realize the role of inequality and power in economic decisions. Individual and firm decisions can be influenced by social and political factors such as conventions, habits, emulation, market structures and even fraud (Galbraith, 1958/1969, 1974, 1996). This can lead to aggregate suboptimal outcomes that do not protect environmental goods and natural capital. There are also inefficiencies associated with prisoner's dilemma, dictator and ultimatum games. Markets by themselves cannot resolve these dilemmas (Kahneman et al., 1986a, 1986b). Given the uniqueness and life-giving importance of natural capital for sustainable development, we cannot leave the protection of these resources in the hands of free markets alone and need a 'strong' definition of sustainability to protect these resources.

Growth versus development

Related to the debate over a 'strong' or 'weak' definition of sustainability is whether we should equate economic growth with economic development. In mainstream economics, growth and development are generally used interchangeably (Brinkman, 1995), with the assumption that economic growth will bring better quality of life and increase the standard of living (Friedman, 2005). If one looks carefully at neoclassical growth models, private manufacturing capital does all the heavy lifting, along with technology. Since economic growth is equated with economic development, public policies for economic development are associated with the expansion of private capital, that is, low capital gains taxes and less environmental regulation on private capital. Private capital is 'the goose that lays the golden egg' for development. Public spending is there to build infrastructure to support private capital, and investments in human capital are evaluated by the quality they add to it. The composition and distribution of economic growth are to be resolved through the market or political process (Greenwood and Holt, 2010a).

The problem with this view is the assumption that economic growth from private manufacturing capital will give us sustainable development

(Earl, 1995). Private manufacturing capital bears the primary responsibility in neoclassical economics for growth, but we should remember that it is derived from natural and human capital (Greenwood and Holt, 2010b, pp. 5–7). Sustainable development requires not relying on one type of capital over another. This means understanding how we can sustain all capital stocks that contribute to our economy, society and environment (Zolatas, 1981; Norgaard, 1994; Daly, 1996; Greenwood and Holt, 2010b; Holt and Greenwood, 2014). By expanding private manufacturing in areas that destroy wetlands, this could lead to flooding due to lack of water absorption. If one recognizes that sustainable development means having an awareness of the interrelationships between various capital stocks, we might be able to develop public policies that avoid this type of problem.

Many non-market aspects that we would want as part of our standard of living – such as health, public safety and climate stability – do not necessarily move in tandem with total output and income. Growth in average gross domestic product (GDP) over the last decades in the United States has not trickled down to provide more well-being for all (Zacharias et al., 2014). Also, economic growth can be volatile and temporary, possibly setting the economy on a path that is economically, environmentally and socially unsustainable (Holt, 2005, 2009). Finally, development should provide improvements in the standard of living that are broadly based. When the benefits of development are skewed toward a particular group, or when the costs of economic growth are borne by certain groups over others, we have not achieved sustainable development. When economic growth leaves too many people behind, it does not improve the overall quality of life. For at least the last 30 years, such inequality has been noticed in the United States, with steady economic growth that has benefitted particular groups over others (see Jones and Weinberg, 2000; Piketty and Saez, 2003; Holt and Greenwood, 2014).

This does not mean that those who hold to sustainable development are anti-growth. Economic growth brings income and consumer spending that truly can increase our standard of living. But economic growth can also take away from our standard of living by creating more congestion, urban sprawl, pollution, inequality, changes in our social and private lives that make civic and personal engagements more difficult, which can possibly lead to a 'negative' trickle-down effect (Greenwood and Holt, 2012). The consequence is that economic growth can either add or take away from our standard of living, and not be sustainable.

Sustainable development differs from neoclassical growth theory in three important ways:

- Instead of having a trickle-down effect, we should have a 'percolating up' model that takes into consideration economic opportunity, environmental and quality-of-life issues for future development. This recognizes that economic growth alone will not solve all the problems associated with sustainable development, and there can be trade-offs between economic growth and economic development.
- There is an appreciation and understanding of the roles of all capital stocks and their contributions to sustainable development.
- We need to go beyond ideology and to be pragmatic in our approach to development. When market forces work, we should use them, but some forms of capital require more public investments and protection than others. An ideology against public investment will not allow us to achieve the level of sustainable development that we need.

Let us now look at the role that heterodox economics has played in helping to define true economic development that benefits the economy, society and environment.

Heterodox economics and true economic development

Institutionalists were the first to make the distinction between economic growth and development, as they saw growth as necessary, but not sufficient for economic development (Galbraith 1958/1969, 1996; Ayres, 1962; Myrdal, 1973). While the neoclassical approach focuses on allocation, institutionalists focus on improving value with development. They look at economic development as an evolutionary process pushed by technology. Once established in a community or society, the overall standard of living should improve for a broad spectrum of the population. Klein (1974, p. 801), representing the institutionalist approach, states about development that: 'the traditional emphasis . . . is on allocation rather than valuation. Progress involves valuation through time, while growth involves simply an increase in whatever it is the economy happens to be doing'. Myrdal (1973, p. 190) expressed the same view in a different way: 'the upward movement of the whole social system . . . not only production, distribution of the produce, and modes of production are involved but also levels of livings, institutions, attitudes, and policies'. Veblen captured the qualitative changes that occur in development as compared to economic growth measured by total output when he stated that true development requires fundamental changes in institutions, how we do things and think (Veblen 1914/1922, 1919/1961; Greenwood and Holt, 2010a, p. 164).

John Kenneth Galbraith brought power into the picture by showing how large corporations and economic interests in an affluent society can set the

agenda for economic growth over development: 'Growth, being a paramount purpose of the society, nothing naturally enough is allowed to stand in its way. That includes its . . . diverse effect, on the environment, on air, water, the tranquillity of urban life, the beauty of the countryside' (Galbraith, 1974, p. 286). He emphasized the same issue in *The Affluent Society* (Galbraith, 1958/1969) when he wrote about 'private wealth and public squalor'. But one of his major contributions to development, which is often overlooked, is that he laid the foundation for a theory of human capabilities, which was later developed by Amartya Sen (1993).

Galbraith was also one of the first economists to show problems with using traditional economic measures of growth as a way of estimating social well-being (see Galbraith 1958/1969). He questioned early on the relation between economic growth and quality of life, and explained why we need a new measurement for economic development that includes not just growth, but also quality of life and sustainability (see Galbraith 1958/1969, 1971, 1996). His pioneer work has led to more recent scholarship in quality of life, sustainable economic development and human capabilities (see Allardt, 1993; Power, 1996; Sen, 1999). This has helped us to redefine social welfare and provide critiques of traditional welfare economics (Sen, 1987; Norgaard, 1994; Power, 1996).

Besides the institutionalists, another important heterodox school to influence the meaning of sustainable development is that of ecological economics. Following the work of Georgescu-Roegen, Boulding and Daly, ecological economists have done significant work showing that natural resources are not 'free gifts from nature', but are resources that can be depleted (see Holt et al., 2009). They have also made it clear that natural capital is more than just an input in the production process, for it includes life-giving elements of the biosphere that have no substitutes (ibid.). Overall, they questioned the feasibility of economic growth over the long run and economic development policy needed for fundamental reform. Some of the primary concerns of ecological economists with undifferentiated economic growth would be the following:

- the size and growth of populations and their impact on economic, social and environmental systems;
- the need to go beyond substitution effects of capital and to recognize the role of different capital stocks;
- endless economic growth is simply unsustainable for both society and the environment;
- the neoclassical assumption that 'more is better' needs to be questioned for both its quantitative and its qualitative meaning;

BOX 20.2

KEYNES, BOULDING AND THE POST-KEYNESIAN SUSTAINABILITY MODEL

Though John Maynard Keynes did not write about the environment or sustainability, he did make a statement about population growth, natural resource capacities and what we should expect in the future. In his *Essays in Persuasion* (Keynes, 1931/1963), he held an optimistic view about economic growth by capital accumulation and technological development. Keynes felt that with the use of the right economic model the world would have adequate resources and be able to solve future economic problems. He did provide the warning that it would depend upon the ability to control the world's population. Keynes's optimism contrasted with John Stuart Mill's pessimism; Mill argued that any type of economic growth, mainly pushed by capital stock, must come to an end. Kenneth Boulding (1966) also took this position when he said that only a fool or an economist would believe that we can have continuous economic growth without damaging the planet with finite resources.

Combining Keynes with Boulding, post-Keynesian economists have developed a sustainable growth model (Holt, 2005, 2007, 2009) that takes into consideration Keynes's principles of radical uncertainty and the irreversibility of historical time with Boulding's (1966) 'Spaceship Earth' model, which includes natural capital and waste. Post-Keynesian environmental economics recognizes that the natural environment, the economy and society are interrelated, with complicated policies being made in a world of uncertainty that includes limited natural resources and waste. There is a debate among environmental post-Keynesian economists over whether we should move from a mechanistic to a biological approach in macroeconomic analysis (Holt et al., 2009). However, in both cases, there is an awareness of the connection between the economy, the environment and society.

- compared to neoclassical economics, we need a dynamic and complex model instead of one that is static.

This leads us to post-Keynesian economists and their work with radical uncertainty (Box 20.2).

While ecological economists have explained the importance of preserving certain types of natural capital that are non-replaceable and life-giving, post-Keynesians have pointed out that the future is unpredictable, based on the relation between physical, biological and institutional factors in the development process. The future is uncertain rather than known, or known with some probability distribution (see Keynes, 1936, 1937). Davidson

(1982–83, 1988, 1991, 1996) has brought this point home by arguing that uncertainty is part of a non-ergodic world, as compared to the ergodic world of neoclassical economics. Systems are ergodic if both their key parameters and their structure are stable over time, which means that we can extrapolate from the past to the future. Non-ergodic systems experience structural change or parameter changes over time, which means that consumers and firms cannot figure out what the future will be like.

Since the consequence of future environmental damage cannot be known, we cannot perform controlled experiments with the Earth and a parallel Earth to see how pollution affects one but not the other. When we deal with economic development that can have a significant impact on the environment, post-Keynesians have argued that we need to recognize the radical uncertainty or non-ergodic world we live in (Holt, 2005, 2009). Yet, we need economic development. Instead of assuming a world where we can calculate and, at some level, predict levels of environmental risk, as in neoclassical environmental management models, the prudent post-Keynesian approach of using complex non-linear dynamic models and applying precautionary principles to derive policies makes more sense. For example, Lavoie (2005) has developed a post-Keynesian theory of consumer choice based on precautionary principles such as procedural rationality, satiable needs, separability, growth and subordination of needs, non-independence and heredity.

Following the work of Lavoie, the second characteristic of post-Keynesian economics that is helpful for understanding the dynamics of sustainable development is that individual decision-making is dependent on social factors, such as human relations, conventions, habits and emulation, rather than on individual rational choice. This means that sustainable development requires us to look at social rationality (as compared to just individual rationality) and at the consequences of these social decisions in development, which can lead us to suboptimal outcomes in markets. The 'prisoner's dilemma' is an example. Such dilemmas, post-Keynesians argue, are common in everyday life, and environmental policy decisions need to take them into consideration. Like the prisoner who confesses, the free-rider in society does not pay to support community needs for clean air and water, because they believe everyone else will contribute to the common good. Of course, if everyone followed this logic, then the aggregate outcomes create serious sustainable problems that affect not just present, but also future generations.

The third insight that post-Keynesians provide for sustainable development is that economic analysis involves an examination of economies moving and evolving through historical time rather than economies that logically adjust

from one equilibrium position to the next. Joan Robinson (1974) criticized neoclassical economics for its preoccupation with equilibrium analysis and the use of logical time. She suggested that economists should focus more on how economies evolve over time, like the institutionalists, instead of how they move to some timeless point. Post-Keynesians, such as Robinson (1974) and Kaldor (1985), believed that economic systems do not have a point of rest and it is better to look at economic development as moving through historical time, where the past and the present matter. By ignoring historical time, neoclassical economists disregard the path dependency of economic systems and their impact on consumers, firms, society and the environment. For example, if we move from one equilibrium point that stresses economic growth and we later become concerned about environmental degradation caused by earlier behaviour, moving to a new equilibrium does not erase that degradation.

The heterodox approach gives us a pluralistic approach to sustainable development. As mentioned, institutionalists were the first to stress the difference between economic growth and development. Ecological economists expressed concerns over the uniqueness of natural capital. Post-Keynesians have shown us the effects of uncertainty, social rationality and path dependency in dealing with sustainable development. We should also mention feminist economics and its focus on 'provisioning' rather than choice as the central problem of economic decision-making, which fits with a sustainable development. As Julie Nelson (1993, p. 296) writes, '[o]ne can think of economics as the study of humans in interaction with the world which supports us – of economics as the study of organization of the processes which provision life'. The heterodox approach gives us a more realistic and richer method to deal with sustainable development by focusing on historical time, radical uncertainty and the uniqueness of natural capital. The heterodox method also does a better job of dealing with power and culture in understanding individual choices and public decisions. As will now be discussed, a pluralistic approach recognizes the interdependence of all capital forms for economic growth rather than attributing almost everything to private manufacturing capital. Finally, and maybe more importantly, heterodox economists recognize the importance of justice, equality and opportunity as part of any true sustainable development model. So what would a sustainable development model look like? How can we develop policies that support and improve sustainable development?

Investments for sustainable development

The primary driving force behind sustainable development is the valuing of all capital stocks used for producing income, quality of life and sustainability.

Besides private capital that yields a profit, there are other types of capital that play roles in economic growth, which include natural, public, social and human capital. All of these are needed for sustainable development. It is important to appreciate the interconnection between these different capital stocks. Instead of looking at private capital as the muscle behind economic development, we now see the interconnection of all forms of capital, with natural capital having unique life-giving qualities that cannot be substituted. This allows us to understand the unintended consequences of the use of one capital stock for another.

For example, as we expand the infrastructure of a town, this might destroy the natural capital of surrounding wetlands. This might create significant flooding due to lack of water absorption, which means building and spending more on infrastructure to solve the flooding problem that was caused by destroying the natural capital. Sustainable development forces us to think about the complex interrelationships between various capital stocks and the costs associated with the development and use of one type of capital over another. Investments in all these capital stocks are there to serve people, reflecting the values and goals of our economy, society and environment. Some of these stocks are owned privately and others publicly, showing the social complexity of their relationship.

Though it might seem odd to think of the environment as an equal partner with private manufacturing capital in development, the following figures might help us to appreciate this relationship. In Figure 20.1 we have the economy, society and environment as separate entities with some overlap.

Figure 20.1 The traditional view of economy, environment and society

The economy is the largest sphere. In this model, manufacturing capital is the foundational force for economic growth and well-being. Other forms of capital such as natural or human capital are secondary, and have value in their ability to support private capital. Natural capital gives us oil, water and land for the production process. Human capital and society are there for quality improvements of private manufacturing capital, to accommodate the needs of private capital for economic growth and to create jobs. Benefits from economic growth are then expected to trickle down to the other kinds of capital (human, natural, social and infrastructure) to sustain them.

Many heterodox economists are working hard to move away from the traditional view of Figure 20.1 to a new paradigm that shows the dynamic relationship between the environment, society and the economy, and the use of different capital stocks (Daly, 1996; Holt et al., 2009; Greenwood and Holt, 2010b). This is represented in Figure 20.2, which tells a very different story from the traditional view.

The economy is now a subset of society and the environment. This allows us to see the synergies between all the capital stocks. It also allows us to see more clearly what are the trade-offs with the use of the different capital stocks. In this model, private capital does not play the major role, as it does in the earlier model for economic development.

Figure 20.2 Sustainable development: economy, environment and society

A model of sustainable development differs in three important ways from the more traditional model represented in Figure 20.1. First, sustainable development does not accept as accurate the 'trickle-down effect', where economic growth based on investment in private capital will sustain all capital stocks needed for the economy, society and nature. Second, all capital stocks need to be protected for their intrinsic value, independently of market values. True development requires investments in all of them, and no capital stock is more important than another. Third, economic growth is not the same as sustainable development measured by output of goods and services. Instead, the goal is improving a broader definition of the standard of living that includes quality of life and sustainability. This gives us a true measurement of economic development that tells us the full cost of development. In the past, the impact of economic growth on the quality of life and the environment has been acknowledged only later in the process of growth. For heterodox economists, these factors need to be recognized upfront, so that we can see the consequences of economic growth on sustainable development.

A sustainable development approach goes beyond our present fixation on private capital. We need to make a distinction between resources that are renewable and those that are not. We need to appreciate where technology or human knowledge can be a substitute for natural capital and where it cannot. Social capital needs to be explored more to understand its positive aspects (see Bowles and Gintis, 2002). Finally, the impact of institutions on development needs to be understood (see Ayres, 1962; Greenwood and Holt, 2008). By recognizing the role of all capital stocks and the unique position that natural capital plays, we are able to have economic development that is based on and around people, that respects our natural resources and the values and goals of social communities. The discussion turns now to creating a better working measurement of economic development.

Measuring a new standard of living for sustainable development

With a new and exciting way to look at development, we need a new measurement. As mentioned, concerns about quality of life, the environment, inequality and future generations have led to widespread dissatisfaction with economic growth as a unilateral goal. This has resulted in efforts to come up with an alternative measurement to GDP. A problem with GDP as a measurement is its inability to differentiate between goods that have a positive effect on sustainability and those that take away from it. GDP measures only production of goods and services that involve a market

transaction, which misses much of the social provisioning that takes place in one's family life and community. For example, a parent staying home to raise children or helping elderly people in one's community can actually lower GDP, though providing very valuable services. States and cities do not have a GDP measure and usually use personal income to measure well-being, but again it only measures economic activity. Providing alternative indicators to at least supplement per capita income or GDP is crucial for us to know what types of policies to put into place for sustainable development, and to make sure that those policies are accountable to the public (Box 20.3).

BOX 20.3

ESTIMATING NON-MARKET VALUES FOR SUSTAINABLE DEVELOPMENT

There are now many new techniques that economists can use to estimate important non-market values for sustainable development. Though mainstream economists do allow estimates of non-market values for ecological services into their cost–benefit models, they are usually limited and restrictive. For example, measuring the damage to external costs is usually associated with a particular type of pollution. Ecological economists have taken on a more ambitious approach of trying to develop a value for the Earth's ecosystem. They do this by evaluating the value of different ecosystem services such as climate control and regulation on different biomes (large natural areas and habitats). Though this method has attracted criticism for trying to reduce the meaning and value of our ecological system to a monetary value, the need to look at our environmental resources as having value is a significant contribution to sustainable development. The value for human well-being from environmental services might be higher than the value they receive from economic growth, which most likely will increase in the future.

The economic discussion of estimating non-market values has led to a debate on whether the focus of environmental value should be anthropocentric. With this view, it is only human worth that counts, and any non-market environmental value must always be judged by the benefit or utility given to humans. We might call this the 'strong' anthropocentric position. Another, which we can call the 'weak' position, recognizes the importance of care and stewardship of all environmental species on Earth, but in an indirect way to support human life. Another view represented by those that support what has been called 'deep ecology' does not necessarily see anything unique about humans as compared to other living species: they should simply be treated as part of nature. The consensus among ecologists is that sustainability ought to be anthropocentric but also respect and protect other species for their intrinsic worth. This is represented by their different methods of measuring non-market values, such as putting a value on environmental qualities or characteristics of nature and species that humans appreciate.

An example of an alternative measurement used for sustainable development is the genuine progress indicator (GPI). The index subtracts some of the costs of economic growth from GDP while adding an estimated value for non-market activities such as household and volunteer work. It is a variant of the Index of Sustainable Economic Welfare (ISEW) first proposed by Daly and Cobb (1994). Developed in 1994, the GPI gives us a better measurement of economic well-being. The measurement particularly addresses depletion of non-renewable resources and the costs of pollution, commuting, crime and other external factors that might affect quality of life and sustainability.

Another example of an alternative measurement comes from the work of Amartya Sen. Traditional welfare economics argues that lowering a person's income will make them worse off, while more income increases well-being. But, as Sen points out, an increased level of consumption of meaningless goods and more production by menial labour (both registered as positive outcomes in traditional measurements of economic growth) can have a negative impact on people's quality of life. It can also fail to provide them with more freedoms, and in fact limit their choices (Sen, 1999). Sen's pioneering work has shown that one can be in a community with lower per capita income as compared to another and be personally much better off if that community provides larger options for personal growth such as nourishment, education, health care and a safe environment; all of which greatly increase the person's welfare more than just a change in their income (see Sen 1987, 1993, 1999). Sen's writings go beyond economic growth and extend to the goals of sustainable development by incorporating human capabilities.

Sen's work on welfare as capabilities has led to a variety of new applications in public policy in improving lives in developing countries. Besides criticizing traditional welfare economics, Sen has developed new measurements for indexing quality of life by looking at poverty, inequality and gender. All of these indexes move beyond income as a way to measure well-being, pushing the notion that we must take things other than income into consideration. Probably the most influential one is the Human Development Index (HDI). Created for the United Nations, the HDI takes into consideration not just per capita income, but also life expectancy, adult literacy, and health. The HDI became part of the United Nations' (1990) first annual *Human Development Report*, when Sen was working there as a consultant. The index is a weighted average of income that is adjusted for distribution and purchasing power and other factors that weigh in human capabilities. The index is classified by the deprivation of what is potentially achievable. The value of HDI ranges from 1 to 0, with 1 being the highest. This measure has encouraged developing countries to look beyond economic growth as a sole judge of policy,

to income distribution, education, health care, life expectancy and safety as important components of social welfare (Greenwood and Holt, 2010a, p. 171).

Public attention to and interest in these alternative measurements is starting to attract the attention of different cities and states throughout the world. The popularity of these new indicators demonstrates two aspects of our changing world. First, higher income and economic growth do not capture many of the social and environmental amenities of quality of life that are important for development. Second, new indicators register increased costs associated with economic growth such as degradation of our environment. To understand what new policies to put in place to support the demand for a better quality of life, and to lower costs associated with economic growth, these new indicators will play a vital role in providing new data on issues surrounding sustainable development.

Conclusion

As we face a new future with new global challenges, the old neoclassical model of economic growth is not up to the job. It is time for heterodox economists, who have been doing exciting new work in the area of sustainable development, to move forward. Galbraith's (1958/1969) point that he made 50 years ago is still meaningful today: we face a problem of social imbalance between private affluence and public squalor. The social imbalance can best be described as one between all the capital stocks for sustainable development. Making a distinction between economic growth and economic development, and appreciating the role of all capital stocks, is a step in the right direction to overcome this imbalance. This chapter provides an alternative paradigm that puts the economy, society and environment on an equal footing, creating a new model for development. As populations increase and there is more concern with quality of life and environmental issues such as global warming, a new way of looking at sustainable development becomes ever more urgent.

REFERENCES

Allardt, E. (1993), 'Having, loving, and being: an alternative to the Swedish model of welfare research', in M. Nussbaum and A. Sen (eds), *The Quality of Life*, Oxford: Oxford University Press, pp. 88–94.

Ayres, C.E. (1962), *The Theory of Economic Progress: A Study of the Fundamental of Economic Development and Cultural Change*, 2nd edition, New York: Schocken Books.

Boulding, K. (1966), 'The economics of the coming Spaceship Earth', in H. Jarrett (ed.),

Environment Quality in a Growing Economy, Baltimore, MD: John Hopkins University Press, pp. 3–14.

Bowles, S. and H. Gintis (2002), 'Social capital and community governance', *Economic Journal*, **112** (483), F419–F436.

Brinkman, R. (1995), 'Economic growth vs. economic development: toward a conception clarification', *Journal of Economic Issues*, **29** (4), 1171–88.

Brundtland, G. (1987), *Our Common Future: Report of the World Commission on Environment and Development*, Oxford: Oxford University Press.

Coase, R. (1960), 'The problem of social cost', *Journal of Law and Economics*, **3** (1), 1–44.

Daly, H.E. (1996), *Beyond Growth: The Economics of Sustainable Development*, Boston, MA: Beacon Press.

Daly, H.E. and J.B. Cobb, Jr (1994), *For the Common Good: Redirecting the Economy toward Community, the Environment, and a Sustainable Future*, 2nd edition, Boston, MA: Beacon Press.

Daly, H.E. and J. Farley (2004), *Ecological Economics: Principles and Applications*, Washington, DC: Island Press.

Davidson, P. (1982–83), 'Rational expectations: a fallacious foundation for studying crucial decision-making processes', *Journal of Post Keynesian Economics*, **5** (2), 182–98.

Davidson, P. (1988), 'A technical definition of uncertainty and the long-run non-neutrality of money', *Cambridge Journal of Economics*, **12** (3), 329–37.

Davidson, P. (1991), 'Is probability theory relevant for uncertainty? A post Keynesian perspective', *Journal of Economic Perspectives*, **5** (1), 129–43.

Davidson, P. (1996), 'Reality and economic theory', *Journal of Post Keynesian Economics*, **18** (4), 479–508.

DeGregori, T. (1974), 'Power and illusion in the marketplace: institutions and technology', *Journal of Economic Issues*, **8** (4), 759–70.

Domar, E. (1946), 'Capital expansion, rate of growth, and employment', *Econometrica*, **14** (2), 137–47.

Earl, P. (1995), *Microeconomics for Business and Marketing*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Friedman, B.M. (2005), *The Moral Consequences of Economic Growth*, New York: Vintage Books.

Galbraith, J.K. (1958/1969), *The Affluent Society*, 2nd edition, Boston, MA: Houghton Mifflin.

Galbraith, J.K. (1971), 'Economics and the quality of life', in *Economics, Peace and Laughter*, Boston, MA: Houghton Mifflin, pp. 3–25.

Galbraith, J.K. (1974), *Economics and the Public Purpose*, Boston, MA: Houghton Mifflin.

Galbraith, J.K. (1996), *The Good Society: The Humane Agenda*, Boston, MA: Houghton Mifflin.

Greenwood, D.T. and R.P.F. Holt (2008), 'Institutional and ecological economics: the role of technology and institutions in economic development', *Journal of Economic Issues*, **42** (2), 445–52.

Greenwood, D.T. and R.P.F. Holt (2010a), 'Growth, development and quality of life', in R. Garnett, E.K. Olsen and M. Starr (eds), *Economic Pluralism*, London: Routledge, pp. 160–75.

Greenwood, D.T. and R.P.F. Holt (2010b), *Local Economic Development in the 21st Century: Quality of Life and Sustainability*, Armonk, NY: M.E. Sharpe.

Greenwood, D.T. and R.P.F. Holt (2012), 'Negative trickle-down and the financial crisis of 2008', *Journal of Economic Issues*, **46** (2), 363–70.

Harrod, R.F. (1939), 'An essay in dynamic theory', *Economic Journal*, **49** (193), 14–33.

Hartwick, J.M. (1977), 'Intergenerational equity and the investing of rents from exhaustible resources', *American Economic Review*, **67** (5), 972–4.

Hartwick, J.M. (1989), *Non-Renewable Resources Extraction Programs and Markets*, Chur, Switzerland: Harwood.

Holt, R.P.F. (2005), 'Post Keynesian economics and sustainable development', *International Journal of Environment, Workplace, and Employment*, **1** (2), 174–86.

Holt, R.P.F. (2007), 'What is post Keynesian economics?', in M. Forstater, G. Mongiovi and S. Pressman (eds), *Post Keynesian Macroeconomics*, London, UK and New York, USA: Routledge, pp. 89–107.

Holt, R.P.F. (2009), 'The relevance of post-Keynesian economics to sustainable development', in P. Lawn (ed.), *Environment and Employment: A Reconciliation*, London, UK and New York, USA: Routledge, pp. 146–60.

Holt, R.P.F. and D.T. Greenwood (eds) (2014), *A Brighter Future: Improving the Standard of Living Now and for the Next Generation*, Armonk, NY: M.E. Sharpe.

Holt, R.P.F., S. Pressman and C.L. Spash (eds) (2009), *Post Keynesian and Ecological Economics: Confronting Environmental Issues*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Jones, A.F. and D.H. Weinberg (2000), 'The changing shape of the nation's income distribution: 1947–1998', US Census Bureau Current Population Reports, No. P60-204.

Kahneman, D., J.L. Knetsch and R.H. Thaler (1986a), 'Fairness as a constant on profit seeking: entitlements in the market', *American Economic Review*, **76** (4), 728–41.

Kahneman, D., J.L. Knetsch and R.H. Thaler (1986b), 'Fairness and the assumption of economics', *Journal of Business*, **59** (4), 285–300.

Kaldor, N. (1985), *Economics Without Equilibrium*, Armonk, NY: M.E. Sharpe.

Keynes, J.M. (1931/1963), 'Economic possibilities for our grandchildren', in *Essays in Persuasion*, New York: W.W. Norton, pp. 358–73.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Keynes, J.M. (1937), 'The general theory of employment', *Quarterly Journal of Economics*, **51** (2), 209–23.

Klein, P.A. (1974), 'Economics: allocation or valuation?', *Journal of Economic Issues*, **8** (4), 785–813.

Lavoie, M. (2005), 'Post-Keynesian consumer choice theory for the economics of sustainable forest management', in S. Kant and R.A. Berry (eds), *Economics, Natural Resources and Sustainability: Economics of Sustainable Forest Management*, Dordrecht: Kluwer, pp. 67–90.

Lucas, R.E. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics*, **22** (1), 3–42.

Maddison, A. (1991), *Dynamic Forces in Capitalist Development*, Oxford: Oxford University Press.

Meadows, D.H., D.L. Randers and W. Behrens III (1972), *The Limits to Growth*, London: Pan Books.

Myrdal, G. (1973), '"Growth" and "development"', in *Against the Stream: Critical Essays in Economics*, New York: Pantheon Books, pp. 182–96.

Nelson, J. (1993), 'Gender and economic ideologies', *Review of Social Economy*, **5**1 (3), 287–301.

Nordhaus, W.D. and J. Tobin (1972), 'Is growth obsolete?', in *Economic Research: Retrospect and Prospect, Volume 5: Economic Growth*, Cambridge, MA: National Bureau of Economic Research, pp. 1–80.

Norgaard, R. (1994), *Development Betrayed: The End of Progress and a Coevolutionary Revisioning of the Future*, London, UK and New York, USA: Routledge.

Pack, H. (1994), 'Endogenous growth theory: intellectual appeal and empirical shortcomings', *Journal of Economic Perspectives*, **8** (1), 55–72.

Panizza, U. (2002), 'Income inequality and economic growth: evidence from American data', *Journal of Economic Growth*, **7** (1), 25–41.

Partridge, M. (2005), 'Does income distribution affect US state economic growth?', *Journal of Regional Science*, **45** (2), 363–94.

Partridge, M. (2006), 'The relationship between inequality and labor market performance: evidence from US States', *Journal of Labor Research*, **27** (1), 1–20.

Pigou, A.C. (1929), *Economics of Welfare*, London: Macmillan.

Piketty, T. and E. Saez (2003), 'Income inequality in the United States, 1913–1998', *Quarterly Journal of Economics*, **158** (1), 1–39.

Power, T. (1996), *Environmental Protection and Economic Well-Being: The Economic Pursuit of Quality*, 2nd edition, Armonk, NY: M.E. Sharpe.

Robinson, J. (1974), *History versus Equilibrium*, London: Thames Polytechnic.

Romer, P. (1986), 'Increasing returns and long-run growth', *Journal of Political Economy*, **94** (5), 1002–37.

Romer, P. (1994), 'The origins of endogenous growth', *Journal of Economic Perspectives*, **8** (1), 3–22.

Sen, A. (1970), *Collective Choice and Social Welfare*, London: North-Holland.

Sen, A. (1981), *Poverty and Famines: An Essay on Entitlement and Deprivation*, Oxford: Oxford University Press.

Sen, A. (1987), *On Ethics and Economics*, Oxford: Basil Blackwell.

Sen, A. (1993), 'Capability and well-being', in M. Nussbaum and A. Sen (eds), *The Quality of Life*, Oxford: Oxford University Press, pp. 30–53.

Sen, A. (1999), *Development as Freedom*, New York: Knopf.

Solow, R. (1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70** (1), 65–94.

Solow, R. (1974), 'The economics of resources or the resources of economics', *American Economic Review*, **64** (2), 1–14.

Solow, R. (1986), 'On the intergenerational allocation of natural resources', *Scandinavian Journal of Economics*, **88** (1), 141–9.

Solow, R. (1992), *An Almost Practical Step Toward Sustainability*, Wallingford, UK and New York, USA: Resources for the Future.

Solow, R. (1994), 'Perspectives on growth theory', *Journal of Economic Perspectives*, **8** (1), 45–54.

Stern, N. (2007), *The Economics of Climate Change: The Stern Review*, Cambridge, UK: Cambridge University Press.

Swan, T.W. (1956), 'Economic growth and capital accumulation', *Economic Record*, **32** (2), 334–61.

United Nations Development Programme (1990), *Human Development Report*, New York, USA and Oxford, UK: Oxford University Press.

Veblen, T. (1914/1922), *The Instinct of Workmanship and the State of the Industrial Arts*, New York: B.W. Huebsch.

Veblen, T. (1919/1961), 'On the nature of capital', in *The Place of Science in Modern Civilization*, New York: Russell & Russell, pp. 324–86.

Zacharias, A., E.N. Wolff and T. Masterson et al. (2014), 'Economic well-being in the United States: a historical and comparative perspective', in R.P.F. Holt and D.T. Greenwood (eds), *A Brighter Future: Improving the Standard of Living Now and for the Next Generation*, Armonk, NY: M.E. Sharpe, pp. 23–50.

Zolatas, X. (1981), *Economic Growth and Declining Social Welfare*, New York: New York University Press.

A PORTRAIT OF AMARTYA SEN (1933–)

Amartya Sen is one of the most creative and respected economists of our time. Besides an impressive academic career, he also served as President of the American Economic Association and received the Nobel Prize for Economic Science. Sen is mostly known for his work in welfare economics and economic development. However, his contributions far exceed any particular field, by broadening our conception of economic well-being and making a significant contribution to some of the basic ideas behind sustainable development.

Sen was born in Maniganj, Bangladesh, on 3 November 1933, to a Bengal family. He was educated at Presidency College in Kolkata, India, where he received his first BA in Economics, later he went to Trinity College, Cambridge, UK, and received his second BA in 1955, and his PhD in 1959. Over his long career he has taught at several universities, including the Universities of Jadavpur and Delhi in India, and the London School of Economics, the University of London and the University of Oxford in the UK. In 1988 he moved to Harvard University, USA, and taught there for ten years. In 1998 he was appointed master of Trinity College, Cambridge, a position he held until 2004 when he returned to Harvard as Lamont University Professor. He won the Nobel Prize in Economics in 1998 for his work in welfare economics.

Sen is best known for his influential book, *Collective Choice and Social Welfare*, published in 1970, where he addressed issues surrounding individual rights and majority rule. Extending his work, he focussed on a new measurement and definition of welfare that is not captured by neoclassical methods, with his Human Development Index (HDI) as an example. Besides his work with welfare economics, Sen is known for his research in economic development, which came about from his personal experience of the Bengal famine in 1943. This experience greatly influenced his second most famous book, *Poverty and Famines: An Essay on Entitlement and Deprivation* in 1981. Sen argued that with the Bengal Famine, the problem was not an adequate food supply, but the distribution of the food to particular groups. This showed that with many famines the cause had nothing to do with food shortages, but involved social and economic factors. Both these books show Sen's lifetime interest in distributional issues and concern for impoverished members of society.

There does seem to be a central theme throughout Sen's work, and that is with human capabilities. We see this in his work in welfare economics, where he questions the neoclassical assumption that individuals are rational utility maximizers. Sen argued that this leaves out too much of what people want and need in life, and makes them look like 'rational fools'. Sen's alternative has been his capabilities approach, which sees welfare economics as developing individual abilities. We also see the capabilities approach in his economic development theories. Overall, Sen's method provides new insights for heterodox economists interested in welfare economics and economic development, and in sustainable development. Contrary to the neoclassical approach that looks at economic growth as being the same as economic development, Sen argues that we must differentiate between growth and development, and recognize the need for sustainable

➡

development to be broadly based and equitable. This requires more attention to the distribution of output, and expanding human capabilities and opportunities in sustainable ways.

Economic development for Sen is about 'expanding the capabilities of people' in a sustainable way that can be measured. He and others have pioneered work at the United Nations in publishing a Human Development Index (HDI) that averages per capita GDP with capabilities variables such as expectancy, adult literacy and sustainability practices. All this fits with the method used in this chapter, with investment in different types of capital stocks to achieve sustainable development. An example Sen gives is of women's rights. By providing women with primary education, health care, safety and self-respect, they will have more opportunities and live better lives. To achieve their full potential and develop their capabilities as human beings, we need to invest and preserve all forms of capital stocks mentioned in this chapter. By maintaining capital stocks such as human, natural, private and social capital for sustainability, and understanding the interaction between these capital stocks, we can provide the investment resources needed for Sen's human capabilities, by providing education through human capital, securing a healthy environment by natural capital, and democratic government with social investments.

Amartya Sen's contributions to economics have been significant and important for those interested in sustainable development. He has insisted that the purpose of economics is more than just producing goods and services, but is also about improving the well-being of people through opportunities to develop their full potential and achieve economic development that is compatible with sustainability. By looking at human economic agents not only as utility maximizers but as individuals who need to develop their human capabilities, and recognizing that development is more than just economic growth, his writings have provided many insights for those interested in sustainable development.

EXAM QUESTIONS

True or false questions

1. Sustainable development is concerned with just the future needs and aspirations of society and not the present.
2. To achieve sustainability, we should just rely on government regulation.
3. Those who support sustainable development look at the human economic system as a subsystem of society and the environment.
4. Economists who are interested in achieving sustainable development rely on just private capital investment to achieve sustainable development.
5. Those who are interested in sustainable development are just interested in the environment, and not in getting rid of poverty and inequality in society.
6. Sustainable development economists do not believe that economic growth by itself will automatically lead to a trickle-down effect of desirable non-market goods such as clean air and water, or always improve quality of life.

7. Because of the economy–environment interdependence, sustainable development economists believe that indiscriminate economic growth should be a long-run goal for all nations.
8. Sustainable development economists believe that other indicators besides per capita GDP should be used to measure well-being.
9. The book *The Limits to Growth*, Meadows et al. (1972) sought to develop a new model that represented essential relationships in the world economy by incorporating more endogenous variables and important relationships between those variables than the traditional neoclassical growth theory model did.
10. The development of new technologies will be enough to achieve sustainable development.

Multiple choice questions

1. The idea of 'meeting the needs of the present without compromising the ability of future generations to meet their own needs' is known as:
   a) economic growth;
   b) third-party effects;
   c) Coase theorem;
   d) sustainable development.
2. Which one of the following statements does not indicate a critical belief in sustainable development?
   a) Sustainable development places greater emphasis on the importance of energy resources.
   b) Sustainable development places greater emphasis on the carrying capacity of the environment.
   c) Sustainable development places greater emphasis on capability development.
   d) Sustainable development places greater emphasis on the flow of goods and services between consumers and firms.
3. Which statement captures an ecological economist's view about market prices in capturing ecological value?
   a) Prices determined by market forces correctly give us the right ecological value.
   b) Market prices always give us prices that are too low.
   c) Market prices always give us prices that are too high.
   d) Prices imperfectly capture the complexity of ecological prices.
4. In developing policies for sustainable development:
   a) all forms of capital stocks must be preserved;
   b) by economic growth alone we can achieve sustainable development;
   c) attempts to value natural resources in economic terms must be avoided;
   d) property rights must be assigned for all natural resources.
5. Which one of the following relationships is not included in the neoclassical circular flow model?
   a) The demand for goods and services from consumers to firms.
   b) The flow of pollution from firms to households.
   c) The flow of money from consumers to firms.
   d) The flow of wages, interest and resource payments from firms to consumers.
6. What was predicted by the basic 'limits to growth' model?
   a) Population growth would increase but then start to decrease in the future to a stabilizing level.
   b) Global warming is caused only by economic growth.

c) International cooperation on environmental issues is all that we need to deal with global environmental issues.
d) Resource shortages and rising pollution would limit economic growth.

7. Which one of the following statements about sustainable development is false?
a) Sustainable development is concerned with environmental and social factors as well as private production resources for development.
b) Sustainable development will provide all humans with what they need without concern with the environment.
c) Sustainable development looks at all capital stocks for development.
d) Sustainable development requires thinking about production and consumption in the present and in the future.

8. Which of the following statements best captures how those that support sustainable development look at the future for the global environment?
a) Ecological collapse is inevitable.
b) Fossil fuels will be required to meet our future energy needs.
c) Economic growth must be balanced with ecological carrying capacity.
d) Birth control strategies must be put into place worldwide.

9. What type of sustainability is generally associated with neoclassical economic theory?
a) No sustainability.
b) Limited sustainability.
c) Weak sustainability.
d) Strong sustainability.

10. Which one of the following statements represents the characteristics of strong sustainability?
a) Strong sustainability assumes that many environmental resources cannot be substituted for other forms of capital.
b) Strong sustainability recognizes the need for all capital stocks and not just material capital.
c) Strong sustainability requires both private and public assets to achieve sustainable development.
d) All of the above.

21

Conclusion: do we need microfoundations for macroeconomics?

John King

OVERVIEW

This chapter:

- explains that the microfoundations dogma threatens the existence of macroeconomics as a separate subdiscipline; it has damaging implications for macroeconomic policy, since it implies support for fiscal austerity, wage cuts and deflation;
- points out that the microfoundations dogma emerged in the 1970s in the context of the 'Great Stagflation';
- shows that the mainstream arguments in favour of microfoundations are not strong; the two most important criticisms involve downward causation and the fallacy of composition;
- points out that the fallacy of composition is involved in some important macroeconomic 'paradoxes';
- suggests some alternative metaphors for the relations between microeconomics and macroeconomics;
- draws some general lessons from the question of microfoundations.

KEYWORDS

- **Metaphor:** It is the application of a name or descriptive phrase to an object or action to which it is not literally applicable, for example 'food for thought' or 'leaving no stone unturned'.
- **Fallacy of composition:** It is the logical fallacy that may arise when we

ignore the possibility that a statement that is true of any individual considered separately is false when applied to them all, taken together.
- **Emergent properties:** Properties of a system that cannot be predicted, even if we have a complete understanding of all the basic features of that system.
- **Downward causation:** Cause–effect relations that run downwards, from macro to micro, for example from society to its individual members, or from the economy to the individual agents who operate in it.
- **Methodology:** The view of reality adopted by practitioners of a scientific discipline and the mode of reasoning that they follow, based on their understanding of ontology and epistemology derived (consciously or not) from the philosophy of science.
- **Micro-reduction:** A special case of inter-theoretic reduction – the attempt to explain one theory in terms of another theory from a different domain – in which macroeconomic problems are solved by reference to microeconomic theory.
- **Microfoundations:** A spatial or constructional metaphor used to justify a micro-reduction strategy, in which the neoclassical model of utility-maximizing individuals underpins the aggregate economy and permits the reduction of macroeconomic theory to microeconomics.

Why are these topics important?

The disappearance of macroeconomics

Since the publication of Keynes's (1936) *General Theory of Employment, Interest and Money*, macroeconomics has been regarded as a separate, semi-autonomous subdiscipline within economics, on a par with microeconomics and every bit as important. The dogma that we must provide microfoundations for macroeconomics, however, threatens to destroy this view, since it involves reducing macroeconomics to microeconomics and doing away with what Keynes described as the 'theory of output and employment *as a whole*' (Keynes, 1936, p. vi).

If the microfoundations dogma is accepted, macroeconomics becomes nothing more than an application of microeconomic theory. There is no longer a case for teaching it as part of the core of an economics degree. It should instead be offered as an option, like industrial organization or environmental economics. This is what the anti-Keynesian, New Classical economist Robert Lucas looked forward to, almost 40 years ago. Before the global financial crisis that erupted in 2007–08, this seemed to be happening.

Implications for economic policy

All this carries some very dangerous implications for macroeconomic policy. In particular, the microfoundations dogma has been used to justify fiscal austerity and wage cuts. The case for fiscal austerity comes from treating the government's finances in the same way as the finances of an individual or a household. In the German Chancellor Angela Merkel's famous words, 'every Swabian housewife knows that you cannot live beyond your means' (Swabia is a region in the South West of Germany, centred on Augsburg). In other words, if households should not live beyond their means, then governments should not, either.

The case for wage cuts to reduce unemployment also seems to be straightforward. In any individual market (for labour or for goods), the existence of excess supply proves that the price (or wage) is too high, and needs to be reduced. So, when there is unemployment, real wages need to be reduced. For the same reason, deflation (a falling price level) seems clearly to be a good thing, since price cuts eliminate excess supply. Wage and price flexibility (downwards) seems to be a good thing, while wage and price rigidity is bad.

However, previous chapters in this volume have suggested that these policy conclusions are quite wrong. First, a government is not a household, and the rules that quite rightly apply to the finances of individuals cannot be applied to public finance without causing serious macroeconomic damage. For the government, the correct principle is 'functional finance', not 'sound finance' (see Chapter 12). Second, the level of employment as a whole is determined in the product market, not in the labour market. It depends on the level of effective demand, which is a macroeconomic variable; easily the most important macroeconomic variable, as shown in Chapter 10. Third, there are very good reasons for preventing deflation, as Keynes (1936) explained in Chapter 19 of *The General Theory*. His conclusions have been vindicated in recent history by the poor performance of the Japanese economy in the three decades since 1990.

The mainstream perspective

A textbook example

One graduate textbook, published in 2008 by the prestigious Princeton University Press, begins by distinguishing 'modern macroeconomics' from the old-fashioned Keynesian variety. In modern macroeconomics, the

author explains, 'the economy is portrayed as a dynamic general equilibrium (DGE) system that reflects the collective decisions of rational individuals over a range of variables that relate to both the present and the future. These individual decisions are then coordinated through markets to produce the macroeconomy' (Wickens, 2008, p. 1). This, the author claims, is a great improvement on 'the traditional Keynesian approach to macroeconomics, which is based on ad hoc theorizing about the relations between macroeconomic aggregates' (ibid.).

This is a rather strange use of the expression 'ad hoc', which the *Concise Oxford Dictionary* defines as 'for this particular purpose, or special(ly)'. Wickens (2008) seems to mean something rather different: 'lacking any foundation in mainstream, neoclassical consumer theory'. This is a criticism often made against the Keynesian consumption function, which links aggregate consumption expenditure to current aggregate income, and is the basis of the income multiplier analysis that underpins the principle of effective demand. Keynesians maintain that these are two separate problems. It is neither necessary nor possible to reduce the macroeconomic issue of aggregate consumption to the microeconomic question of individual consumer behaviour (see Chapter 10). Why should anyone think otherwise?

The origins of the microfoundations dogma

We have to begin a century and a half ago, with the Marginalist Revolution in economics that replaced the classical economists' emphasis on social classes and aggregate economic outcomes with a focus on individual utility maximization. Adam Smith (in 1776) and David Ricardo (in 1817) had analysed the relationship between economic growth and the distribution of the national income between workers, capitalists and landlords. Writing in the 1870s, William Stanley Jevons, Carl Menger and Léon Walras were much more interested in individual decision-making and the coordination of their decisions in competitive markets (see Chapter 3).

This gave rise to the principle of methodological individualism, which was also advocated in the mid-twentieth century by some philosophers of science and social theorists (see Kincaid, 1998). Davis (2003, p. 36) describes it as 'the view that not only can one always in principle replace explanations of social entities by some individualist explanation, but one ought to do so whenever practically possible'. There was a vigorous debate on methodological individualism among sociologists, anthropologists and philosophers of science in the 1950s. It failed to win over most social scientists. The philosopher Steven Lukes (1968, p. 125) dismissed it as 'a futile linguistic

BOX 21.1

METHODOLOGICAL INDIVIDUALISM

Long before there was any talk of micro-foundations, very similar questions had been raised by the advocates of methodological individualism, who also insisted on the need for the reduction of all statements about society to statements about individuals. The expression 'methodological individualism' was invented in 1908 by the eminent historian of economic thought Joseph Schumpeter, but the arguments date back at least to the early marginal utility theorists of the 1870s. They were given added force in the Cold War, when methodological individualism was seen as an antidote to the collectivism and potentially totalitarian 'holism' that opponents of Stalinism attributed to Hegel and Marx and, more legitimately, to some of Marx's followers.

In the 1950s there was a vigorous debate on the issues raised by the defenders of methodological individualism, who supposedly included Karl Popper (though he was not entirely consistent in his support). One of Popper's disciples, and a colleague at the London School of Economics, J.W.N. Watkins, maintained that knowledge of social situations was always obtained from information about individuals, and had therefore to be seen as derived from such knowledge. His critics raised the obvious objection that individual behaviour always takes place within economic and social institutions, the existence of which is already assumed when the study of individuals is undertaken. As the (regrettably non-feminist) philosopher Ernest Gellner put it in 1956: 'History is *about* chaps. It does not follow that its explanations are always in terms of chaps' (Gellner, 1956/2003, p. 14). Thus the critics of methodological individualism were pointing to the fallacy of composition and to the pervasive nature of downward causation, which rendered the case for micro-reduction unconvincing, not just in economics but in all the social sciences.

purism'. 'Why should we be compelled to talk about the tribesman but not the tribe', he asked, 'the bank-teller but not the bank?' (ibid.) (Box 21.1).

For the most part, mainstream economists kept out of this debate. They would not have found it easy to justify reducing the analysis of the banking system to statements about the behaviour of bank tellers. But after 1936 there was always going to be a potential problem for economics in reconciling the new Keynesian macroeconomics, which claimed to provide a distinctive new 'theory of output and employment as a whole', with the old marginalist theory of individual outputs and levels of employment in particular product and labour markets. In particular, there was a potentially serious problem in reconciling the new theory of aggregate employment with the old microeconomics of individual labour markets, broken down by industry, region and level of skill.

In 1956, with the publication of Robert Solow's neoclassical growth model, a new problem was created. There was now an apparent inconsistency between the short-run Keynesian model, in which investment determined saving, and both capital and labour could be in excess supply, and the long-run neoclassical model, in which saving determined investment, and capital and labour were always fully employed. There were two ways out of this contradiction: either apply the principle of effective demand to the long-run model of economic growth, or eliminate it from the short-run model of income and employment. One led to the post-Keynesian analysis of demand-determined growth, the other to the 'real business cycle' models of New Classical economists (see Chapters 10, 13 and 19).

Things came to a head in the crisis years of the 1970s, the era of the 'Great Stagflation', when the Phillips curve became unstable and both inflation and unemployment increased rapidly. Macroeconomic theory seemed to have broken down, and Keynesian macroeconomics in particular was increasingly discredited (as shown in Chapter 10, however, it was only the neoclassical interpretation of Keynes that had failed). There was a particularly important question that no one seemed able to answer: what determined people's expectations of future inflation?

It was vital to have a convincing answer to this question, since inflationary expectations were (and are) an extremely important influence on the actual inflation rate (see Chapter 11). And there was a more general problem, as Lucas pointed out at the time. To predict the consequences of changes in macroeconomic policy, econometricians used forecasting models that relied on the stability of the macroeconomic equations governing the key variables: the consumption function, the investment function, the wage level, the price level. But the parameters of these equations might well be altered by the policy changes themselves, making the predictions unreliable (Box 21.2).

Take the (important) example of inflationary expectations. If firms and households reacted to changes in monetary, fiscal or labour market policy by revising their expectations of future inflation – always in an upward direction, or so it seemed in the 1970s – the forecasting equations would yield the wrong estimates of inflation. Thus successful forecasting requires that the econometricians have some reliable information about how individuals and firms actually form their expectations of future inflation.

No one – least of all Keynes – would deny that microeconomics knowledge can make a significant contribution to macroeconomics. But Lucas went much further, denying the need for a distinct, semi-autonomous subdisci-

BOX 21.2

ROBERT LUCAS

Generally accepted as the founder of the New Classical school of macroeconomics, Robert E. Lucas Jr was born on 15 September 1937 in Yakima, Washington, USA, and grew up in Seattle. He graduated in history in 1959 and earned a PhD in economics five years later, both at the University of Chicago. He taught at Carnegie Mellon University until 1975, when he returned to Chicago, where he is now the John Dewey Distinguished Service Professor Emeritus of Economics. Lucas was awarded the 1995 Nobel Prize in Economics 'for having developed and applied the hypothesis of rational expectations, and thereby having transformed macroeconomic analysis and deepened our understanding of economic policy'. In a 2009 article in the *Wall Street Journal*, N. Gregory Mankiw described Lucas as the most influential macroeconomist of the last quarter of the twentieth century.

In macroeconomic theory, Lucas used the rational expectations hypothesis that he learned from his Carnegie Mellon colleague John Muth to argue for the neutrality of money and the restoration of Say's Law (hence the characterization of his approach as 'New Classical' economics). In terms of macroeconomic policy, Lucas advocated the use of monetary policy to target inflation, and denied the effectiveness of fiscal policy, which he deemed unnecessary for the achievement of full employment; thus he was quite explicitly anti-Keynesian. On the important question of economic methodology, Lucas was an early and extremely influential advocate of the need to provide microfoundations for macroeconomics. Politically, he was and remains a libertarian, providing very clear evidence of the strong connections between free market liberalism and New Classical theory.

pline of macroeconomics altogether. In his New Classical economics, the principles of individual maximization and market-clearing equilibrium are applied directly to all questions of macroeconomic theory and policy: there are no specifically macroeconomic problems to be solved.

Not surprisingly, there was a political element in all of this. It was bound up with the rise of neoliberalism: the doctrine that all social problems have a market solution, and the application of this doctrine to all areas of social and economic policy. Neoliberal politics began with the governments of United Kingdom Prime Minister Margaret Thatcher and United States President Ronald Reagan in the 1980s and is still going strong. Privatization and deregulation were only part of the neoliberal policy package. 'Sound public finance' and price and wage flexibility (downwards) were also called for, or so it seemed to neoliberal macroeconomists, who had rejected the principle of effective demand in the course of their critique of the traditional 'ad hoc' Keynesian theorizing (Box 21.3).

BOX 21.3

NEOLIBERALISM

In Chapter 18, 'neoliberalism' has been defined as the doctrine that markets should dominate the economy and the government should play a relatively small role within the economic system, and its close connection with financialization has been emphasized. It is important to be clear that neoliberalism and the case for microfoundations are logically distinct, in the sense that accepting one of them does not require you to accept the other. A belief in the need to supply microfoundations for macroeconomics does not entail support for neoliberalism, as there may be many substantial market failures at the microeconomic and mesoeconomic levels that require a significant degree of state intervention to put things right. Conversely, support for neoliberalism does not entail acceptance of the microfoundations delusion, at least not in principle.

But in practice the connections are quite strong. Belief in the generally optimal nature of the outcomes that are generated by unregulated markets sits very comfortably with the notion that there are no significant macroeconomic problems with the operation of the capitalist market system. Conversely, support for a broadly Keynesian (and, *a fortiori*, post-Keynesian) approach to macroeconomics is pretty well impossible to reconcile with neoliberalism. Deficient aggregate demand and large-scale involuntary unemployment require government intervention, on a much larger scale than anything the neoliberals would be prepared to tolerate.

The upshot was an insistence on the use of models with RARE microfoundations: that is to say, models that began by assuming the existence of representative agents with rational expectations (RARE). Both halves of this set of initials are problematic. The good reasons for objecting to rational expectations have already been set out in Chapter 8. The assumption of representative agents is even less plausible, since it entails that all individuals are identical. There is thus no reason for them to trade with each other, no reason why their decisions should be coordinated, and therefore no role for markets.

Some mainstream arguments in favour of microfoundations

Quite often, no explicit reason is offered for the insistence on providing RARE microfoundations for macroeconomics, or indeed any other sort: it is simply taken for granted. After all, who could possibly complain about a building having foundations? In fact, this is a good example of question-begging or persuasive language, like 'free market economics' (in Australian English, 'economic rationalism'). Who would want to be unfree (or irrational)? Again, the political dimension is lurking just below the surface. The

philosopher George Lakoff (1996) has shown how the conservative Right has succeeded in cornering the market in persuasive metaphors – 'moral strength', 'moral bounds', 'moral nurturance', 'the nation as family' – leaving the liberal and social democratic Left floundering. The 'microfoundations' metaphor has achieved something similar for neoliberal economics and thus (indirectly) for neoliberal politics.

Where mainstream economists do try to justify the microfoundations metaphor, they use arguments such as the four that follow. First, microeconomics is more basic than macroeconomics, so that in order to be persuasive, sound, reliable and robust, macroeconomics must make explicit reference to microeconomics. Second, the economy is made up of individuals, who must therefore be the starting point for any analysis of the ways in which they interact. Third, microeconomic models have been used successfully to deal with all manner of social and political questions, such as voting behaviour, crime, discrimination and the family, so they should also be applied to macroeconomics. Fourth, micro-reduction has succeeded in the natural sciences, most obviously in biology with the triumph of modern genetics, and the principle should therefore be extended to economics (Box 21.4).

BOX 21.4

REPRESENTATIVE AGENTS

There was always something distinctly odd about the 'RA' component of the acronym 'RARE', which was used by proponents of the New Classical macroeconomics from the early 1970s onwards. Whatever your attitude to the idea of rational expectations (the 'RE' component), it would be difficult not to be unhappy with the assumption of a single representative agent. If all the individual actors in the economy under consideration were identical, it is hard to see why they would trade with each other, and even harder to accept that any such RARE model could be seen as having provided any sort of microfoundation for macroeconomics.

Many mainstream macroeconomic theorists have been uncomfortable with the use of the representative agent assumption, and a variety of attempts have been made to replace it with something more sensible. Whether the use of two or more diverse agents really solves the underlying problem remains contentious. Most recently, some heterodox economists, as well as some on the fringes of the mainstream, have started to develop agent-based models, in which a large number of heterogeneous individuals interact to generate interesting and often quite complex macroeconomic outcomes. Whether these models represent an abandonment of the microfoundations dogma, or simply reintroduce it in a less obviously objectionable but possibly more insidious form, also remains to be seen.

This final argument is worth pursuing. Richard Dawkins, best-selling author of *The Selfish Gene* (1976), advocates the general principle of hierarchical reduction. In another book, *The Blind Watchmaker*, he uses a revealing example:

> The behaviour of a motor car is explained in terms of cylinders, carburettors and sparking plugs. It is true that each of these components is nested atop a pyramid of explanations at lower levels. But if you asked me how a motor car worked you would think me somewhat pompous if I answered in terms of Newton's laws and the laws of thermodynamics, and downright obscurantist if I answered in terms of fundamental particles. It is doubtless true that at bottom the behaviour of a motor car is to be explained in terms of interactions between fundamental particles. But it is much more useful to explain it in terms of interactions between pistons, cylinders and sparking plugs. (Dawkins, 1996, p. 12)

The hierarchical reductionist, Dawkins continues:

> explains a complex entity at any particular level in the hierarchy of organization, in terms of entities only one level down the hierarchy, entities which, themselves, are likely to be complex enough to need further reducing to their own constituent parts, and so on . . . the hierarchical reductionist believes that carburettors are explained in terms of smaller units . . . which are explained in terms of smaller units . . ., which are ultimately explained in terms of the smallest of fundamental particles. Reductionism, in this sense, is just another name for an honest desire to understand how things work. (Dawkins, 1996, p. 13)

Perhaps the provision of microfoundations for macroeconomics is also simply the result of 'an honest desire to understand how things work'? (See Box 21.5.)

A heterodox critique

In fact, none of these arguments is very convincing; Dawkins's hierarchical reductionism least of all, as we shall see shortly. What of the other arguments for microfoundations? What can we learn from the statement that microeconomics is 'more basic' than macroeconomics? The *Concise Oxford Dictionary* offers two definitions of the word 'basic'. The first is 'fundamental', 'serving as the base or foundation'. This provides some synonyms for 'basic', but otherwise does not take us any further. It is just a different way of stating the same thing. The second definition is 'constituting a minimum', and an example is provided: Eleanor Roosevelt's (1961) description of a poor Asian country where 'the basic food of the people is rice for every meal'. Other examples given are 'basic English' (with a very limited vocabulary) or 'basic pay' (the

BOX 21.5

RICHARD DAWKINS

Richard Dawkins was born in the then British colony of Kenya in 1941, but moved to England with his family in 1949. Educated at Oundle College and Balliol College, Oxford, he first studied and then taught zoology until, in 1995, he became Oxford's first Professor for the Public Understanding of Science. Dawkins is best known as a highly articulate, outspoken and forthright advocate of one particular, highly controversial, micro-reductionist version of the theory of evolution. He also won a very large readership, and achieved a degree of notoriety, for his attack on mainstream religion in his book *The God Delusion* (Dawkins, 2006).

Not surprisingly, Dawkins has made some high-profile enemies. In his autobiography, *Candle in the Dark*, he responds to some of them, and also acknowledges his intellectual supporters, some of whom regard him as being an eminent philosopher of science in addition to a practising scientist (Dawkins, 2015, pp. 95–6, 333–4). He has certainly raised some important philosophical issues, one of them being the use of metaphors in scientific writing. 'The selfish gene', which formed the title of an early best-selling book, is a good example: in its literal sense selfishness is of course a property of creatures, above all human creatures, not of miniscule components of their bodies. Similarly with the notion of 'evolutionary stable strategies', since in everyday English strategies are employed by senior army officers, not by animals or their genes. Dawkins also refers approvingly to the 'evolutionary game theory' advocated by his friend John Maynard Smith, again without noting that games are played by humans (together with many species of animals and birds), not in any literal sense by genes. The appropriateness of metaphors remains a contentious question, in biology no less than in economics, and the use of economic metaphors in biology is especially intriguing.

lowest possible wage; in Australia the legal minimum wage was for many years known as the 'basic wage'). This does not seem to help the case for microfoundations any more than the final example that the dictionary offers: 'basic industry', meaning 'an industry of great economic importance'.

What of the argument that, since 'the economy is made up of individuals', we have to start by analysing individual behaviour? This involves a rather elementary confusion between ontological reduction and explanatory reduction. To say that *A* (the economy) is made up of *B*s (individual human beings) is to say something about ontology: what exists. It is rather obviously true. But it does not follow from this that *A* can and must be explained in terms of statements about *B*s, and only in terms of statements about *B*s. We have already encountered one example, with *A* as the banking system and *B*s as the tellers (and other employees) who work for the banks.

Consider a more personal example. There is a sense in which you and I are made up of our body parts: head, toes, arms, legs, vital organs, naughty bits. But we would feel insulted if someone claimed to be able to explain our character, personality, thoughts, emotions and behaviour solely by reference to those body parts. Something – probably a great deal – would be missing from such an explanation.

The third argument relies on the apparent success of the 'economics imperialism' project. Mainstream microeconomic models and econometric techniques have – supposedly – been applied to problems that were previously the preserve of political scientists, sociologists, anthropologists and other lesser breeds, with such success that they have been forced to give up much of their intellectual territory to the economists. You wish! This is what leading economic imperialists such as Gary Becker would like to have happened, but there is not much evidence to show that they have succeeded. It is not a strong argument for the application of micro-reduction to macroeconomics (Box 21.6).

Finally, there is the argument that hierarchical reduction has succeeded in the natural sciences, above all in biology, and should therefore be applied to economics. Many biologists agree with Dawkins, but many do not. In a famous critique, *Not in Our Genes*, Richard Lewontin, Steven Rose and Leon Kamin emphasized the complexity of the relationship between an organism and its environment:

> Organisms do not simply adapt to previously existing, autonomous environments; they create, destroy, modify, and internally transform aspects of the external world by their own life activities to make this environment. Just as there is no organism without an environment, so there is no environment without an organism. Neither organism nor environment is a closed system; each is open to the other. (Lewontin et al., 1984, p. 273)

This points to two insuperable difficulties with Dawkins's argument: downward causation and the fallacy of composition.

Downward causation and the fallacy of composition

The difficulties posed for Dawkins by downward causation can be seen very clearly from his automotive example. Changes in the social, political, economic and cultural context in which cars are driven frequently affect not just the car as a whole machine, but also some or all of its parts. Thus causation runs downwards, from the larger to the smaller units, and not just upwards from the smaller to the larger, as Dawkins maintains. To understand the

BOX 21.6

ECONOMICS IMPERIALISM

The claim that mainstream microeconomics is successfully invading the other social sciences, replacing both their theoretical systems and their methods of empirical research, was always rather fanciful. In fact a reverse process is under way, in which political scientists and sociologists are discovering heterodox economic ideas and finding them to be much more congenial than mainstream thinking. Increasingly, too, heterodox economists are finding jobs outside economics departments and working with colleagues in these related social science disciplines. This seems to be an international phenomenon. In the United Kingdom (UK), the Austrian post-Keynesian Engelbert Stockhammer recently became Professor of Political Economy in the prestigious Kings College London; but in the Department of European and International Studies, not in Economics. Back in Austria, a loose interdisciplinary alliance of social scientists from various disciplines, including some heterodox economists, has intervened for three decades in political debates on such issues as budgetary policy, the distribution of income and the economic implications of European Union enlargement, under the title 'Advice on Social, Economic and Environmental Alternatives', known from its German initials as BEIGEWUM (www.beigewum.at).

This does not represent some insidious form of reverse imperialism, but is evidence of fruitful interdisciplinary cooperation. It has been endorsed by none other than Thomas Piketty, whose best-selling volume *Capital in the Twenty-First Century* (2013) was criticized – correctly – by some reviewers for making too much use of mainstream economic ideas. At the very end of his latest massive book, *Capital and Ideology* (first published in French as *Capital et idéologie*), Piketty writes as follows:

> I am convinced that part of our current political confusion results from the excessive autonomy of economic knowledge vis-à-vis the other social sciences and in the civic and political arena . . . only the interaction of the economic, historical, sociological, cultural and political approaches will allow us to make progress in our understanding of socio-economic phenomena. (Piketty, 2019, p. 1197; our translation)

Amen to that.

causes of changes in car components over time, we need more than knowledge of metallurgy, chemistry and particle physics; more even than complete knowledge of these things. We also need to know about society, politics, psychology and the economy.

There are two well-known examples of downward causation from the mid-twentieth century motor industry. One is provided by the fins and other ornamental embellishments, with no engineering advantages, that were added to cars in the 1950s on the insistence of marketing specialists who drew on

studies of consumer psychology in order to suggest ways of increasing brand loyalty and therefore also profitability. The other was the belated introduction, in the 1960s and 1970s, of a range of safety features in the wake of an intense political controversy over the dangers involved in driving cars without them. In both cases causation ran downwards, from human society to the component parts of its machines; from larger units to smaller units, from macro to micro.

Neither is it possible to infer all the properties of the car from a complete knowledge of its parts. Cars are not just pieces of machinery. They have social, political, economic and cultural significance. They are studied by traffic engineers, by urban sociologists, by town planners and even by political economists who are interested in the demise of the so-called Fordist stage of capitalist development. None of them could conceivably be satisfied with information about car components, not even if it was accompanied by knowledge of metallurgy, chemistry and particle physics. This is not to say that information about the parts might not be useful to them, in some circumstances, or that they should dismiss such information as trivial or misleading. But it would certainly not be sufficient, and would probably not be very enlightening for their particular purposes.

To deny this involves a fallacy of composition: a statement that is true of any individual considered separately may be false when applied to them all taken together. To take an example from everyday life: with the introduction of all-seater football stadiums in the United Kingdom after the 1989 Hillsborough disaster, any individual supporter who stands up to get a better view does indeed get one, but if they all stand up then no one's view is improved.

The fallacy of composition is closely related to the concept of emergent properties, which are defined as properties of a system that are autonomous from the more basic phenomena that give rise to them. In other words:

> a state or other feature of a system is emergent if it is impossible to predict the existence of that feature on the basis of a complete theory of basic phenomena in the system . . . A closely related idea is that emergent properties cannot be explained given a complete understanding of more basic phenomena. (Bedau and Humphreys, 2008, p. 10)

Again, the use of the word 'basic' does nothing to advance the case for the microfoundations dogma.

To return to Dawkins's example: a complete theory of car components would not allow us to explain the social, political, economic and cul-

tural significance of the motor car. These are emergent properties, which cannot be inferred from a theory of the parts and cannot be reduced to it. That is why we need traffic engineers, urban sociologists, town planners and political economists, in addition to metallurgists and motor mechanics.

Some economic examples

In Keynesian macroeconomics there are several well-known cases in which a fallacy of composition can be identified. For some reason they are usually described as 'paradoxes'. First, and most obvious, is the paradox of thrift (see Chapter 10). A decision by any individual to save a larger proportion of their income may lead to more saving by that individual. However, in the absence of increased investment this will not be true of an increase in everyone's savings propensity, which will simply reduce their incomes and leave the volume of aggregate saving unchanged. This proposition is at least 300 years old. As Keynes noted, it was popularized by Bernard Mandeville in his 1714 *Fable of the Bees.*

A realistic monetary theory must allow for the paradox of liquidity. This expression is of quite recent origin, but the principle is long-established in the literature on financial panics (see Chapters 5, 6, 18 and 19). If any individual bank or other financial company wishes to increase its liquidity, it can always do so (at a price). But if all financial companies attempt to do so, the consequence will be a reduction in aggregate liquidity and (in the absence of government intervention) the real possibility that the whole system will collapse. The global financial crisis that began in September 2008 provided a dramatic example of this principle.

The Kaleckian paradox of costs is rather similar. A wage rise is very bad news for any individual capitalist. But it may be good news for them all, taken together, if the consequent rise in consumption expenditure raises the level of economic activity and thereby increases aggregate profits. In practice, this may or may not be the outcome; it all depends on the values of the relevant parameters. But it cannot be ruled out in principle.

Kalecki's profit equation offers yet another example. In a closed economy with no government, and on the assumption that workers do not save, aggregate profits are equal to and determined by the sum of capitalist expenditure on consumption and investment. Thus 'capitalists get what they spend', but only as a class. Any capitalist who thinks that it applies to them as an individual will end up in jail (if they are caught) (Box 21.7).

BOX 21.7

MICHAŁ KALECKI

The great Polish economist Michał Kalecki anticipated some of Keynes's most important theoretical innovations, drawing on Marxian ideas as he did so (see the Portrait of Kalecki in Chapter 12). He wrote little or nothing explicitly on macroeconomic methodology, but often practised it quite brilliantly. Kalecki developed a distinctive non-neoclassical microeconomic theory in which the pricing and investment behaviour of capitalist firms both influences and – crucially – is influenced by macroeconomic developments. In effect he is identifying the horizontal relationship between microeconomic and macroeconomic theory that has been advocated in the present chapter, without ever using the term.

As we have seen in this chapter, Kalecki developed a theoretical model of one very important case of downward causation (that is, causal influence that flows from macro to micro), which has – rather unfortunately – come to be described as a 'paradox'. In the so-called 'paradox of thrift', the decisions of individuals to save more have macroeconomic consequences that negate those decisions. It is a 'paradox' only from the viewpoint of a believer in the need for microfoundations. Non-believers, such as Kalecki, see it as a plausible and potentially serious occurrence in the economic world in which we actually live.

These are all cases where individual behaviour is governed by macroeconomic requirements. The logic of our macroeconomic analysis tells us that, in aggregate, saving cannot increase unless investment rises. The viability of the entire financial system may be threatened by individual firms' quest for increased liquidity. Under some circumstances increased wages may lead to higher profits, not lower profits, in aggregate. Total profits always depend on total spending. None of these results is immediately obvious, none could be inferred from knowledge of microeconomics alone, and every one of them also entails the existence of downward causation, from macroeconomics to microeconomics.

Macrofoundations for microeconomics?

The potential significance of downward causation has led some heterodox economists to argue that we need macrofoundations for microeconomics. There are good reasons for taking this argument seriously: first, because of the irreducibly macroeconomic nature of many important problems (as noted above); and second, because what may be termed the prevailing macroeconomic regime has profound implications for individual behaviour. To understand individual decisions, it is often necessary to know something

about the macroeconomic context in which these microeconomic decisions are being made. The early Keynesians believed that a 'full-employment economy' would be quite different from an 'unemployment economy', especially in the labour market, where – in Kalecki's famous phrase – full employment posed a potentially very serious threat to 'discipline in the factories'.

So the macroeconomic context is very important indeed. Whether it establishes a case for macrofoundations is another matter. There is something perverse about the idea of foundations that exist at a higher level than the edifice that they are supposed to be foundations of, whether the edifice is physical or intellectual in nature. For this reason, let us avoid the term 'macrofoundations'.

A strong case can be made, however, that economists need to provide social and philosophical foundations for their theories, micro and macro (see Chapters 1 and 3). Both macroeconomists and microeconomists need to be aware that they are attempting to model capitalism, not peasant agriculture or handicraft production. There are (at least) two classes of agents – capitalists and workers – and it is the former who own the means of production and control the production and sale of commodities. Firms are not simply the agents of households. Production is motivated by profit, not (at least, not directly) by the utility functions of asocial, classless consumers. Since profit is by definition the difference between revenue and costs – that is, the difference between two sums of money – it is pointless to model a capitalist economy in terms of barter (see Chapter 4).

These social foundations of any meaningful economic theory are inescapable, but they are routinely violated in the mainstream models that employ RARE microfoundations, as can be seen from the opening pages of Michael Wickens's textbook quoted at the beginning of this chapter.

As for the philosophical foundations of economics, a minimum requirement would be some form of scientific realism, including (but not restricted to) a substantial degree of 'realisticness' (see Chapter 1). As Keynes (1936, p. 3) insisted, economic theory should have some bearing on 'the economic society in which we actually live' (Box 21.8).

Some alternative metaphors

We have seen how important it is for economists to be careful with the language that they use, and especially with their metaphors. 'Microfoundations' is a good example of a very bad metaphor. But this does not mean that

BOX 21.8

SCIENTIFIC REALISM

The advocates of scientific realism argue that economic models must display a substantial degree of 'realisticness', by which they mean that these models should be clearly recognizable as simplified versions of a very complex real world. Of course, this is not a black-and-white, either/or condition, as the required degree of realism may well depend on the context, and different observers may well differ in their judgement as to whether the requirement has been met in any particular case. Realists, however, would agree that economic models should be iconic in nature; that is to say, they should resemble maps of towns, which are always much smaller than the areas that they represent but also retain all their most important features. Adjusting the metaphor slightly, to allow for three dimensions rather than two, realists maintain that economic models should resemble the city models that can often be found in local history museums, not the fantastic futuristic constructions that can be found in science fiction movies (beginning with Fritz Lang's classic 1925 film *Metropolis*).

Many heterodox economists have been attracted to one particular version of scientific realism, known as critical realism. Originating with the philosopher Roy Bhaskar, and promoted in economics by the Cambridge economist Tony Lawson, critical realism involves 'open-system' rather than 'closed-system' thinking, so that account can always be taken of new conditions and new variables. Thus its practitioners are highly critical of the deductive modes of reasoning that are employed by many mainstream economists, prefer biological to constructional metaphors, and emphasize the importance of downward causation and emergent properties. However, the restrictions that open-system thinking appears to impose on formal modelling and the use of sophisticated econometric techniques has made many heterodox macroeconomists wary of critical realism and more sympathetic to the less demanding versions of scientific realism proposed by philosophers such as Uskali Mäki and John Searle.

economists should avoid the use of metaphors altogether. Even if this were possible, it would be unwise. In macroeconomics, some hydraulic metaphors are legitimate and helpful: the circular flow of income and expenditure is the most obvious example, along with the circulation of money, and the treatment of investment and saving as injections and leakages, respectively.

Even spatial metaphors have their uses. The great Austrian-American economist Joseph Alois Schumpeter (1883–1950) is said to have described static and dynamic economic theory as being 'separate buildings'; though he did not state which was the *Kindergarten* and which the Institute of Advanced Study. Perhaps the relationship between microeconomics and macroeconomics should also be thought of as a horizontal rather than a vertical one?

I can offer another personal example, this time from my former employer. Many years (and several administrative restructurings) ago, La Trobe University had an Economics building and a Social Science building, side by side, each four storeys high. Between the two buildings there was a bridge at the second level, making it easy to get from the one to the other without going right down to the ground floor. Each building had its own solid foundations but – rather obviously – neither building was the foundation of the other. Occasionally economists and social scientists moved from one building to the other to talk to each other (not often enough, I fear). Perhaps microeconomics and macroeconomics should be thought of in this way? They too are separate, but close to each other, and there is a considerable volume of two-way traffic between them.

One problem with this metaphor is that we probably need more than two buildings, and the network of bridges between them might get very complicated. Between the individual human agent (the 'micro' building) and the global economy (the 'macro' building) there are several intermediate dimensions, including industries, sectors, regions and nations. Schumpeter's modern disciples often refer to the need for 'mesoeconomics' as a sort of middle level of theorizing, between macroeconomics and microeconomics, but even this does not seem adequate.

Just possibly, a Russian doll analogy might be more helpful. There can in principle be any number of dolls. Each one is nested inside the larger one above it. They are similar, but not identical, since the amount of detail has to be reduced as we move from larger to ever-smaller dolls. The largest of the dolls is the most beautiful and the most interesting, but it is also the most complicated and most difficult to understand. The smallest doll is the plainest and least fascinating, but also the easiest to make sense of. There is no case for discarding the largest doll or for concentrating our attention on the smallest.

Of course, this metaphor also has its problems. A set of Russian dolls does not change much over time, give or take some fading or cracking of the paintwork. And they do not interact with each other in any meaningful way. No analogy is perfect, and it is a good idea to bear this in mind when thinking about the use of metaphors by economists. We cannot live without them, and it would be silly to try, but we do need to be careful with them.

Heterodox microfoundations for heterodox macroeconomics?

Unfortunately, some heterodox economists have disregarded this important lesson. It is undoubtedly true that heterodox microeconomics is very

different from mainstream microeconomics (see the Introduction to this book). It emphasizes oligopoly rather than perfect competition, and mark-up pricing rather than Michael Wickens's Walrasian auctioneer. Preferences are viewed as socially determined or endogenous to the economic system: this is another important example of downward causation.

These important differences have led some heterodox economists to claim that they, too, are providing microfoundations for macroeconomics. It is just that theirs are very different, and much better, than the microfoundations supplied by the mainstream. This is an understandable reaction to a very difficult situation. It is hard enough these days for heterodox economists to be taken seriously by the mainstream, and a failure to provide microfoundations seems to make it even harder. But that is still a serious mistake. No heterodox economist would defend the reduction of macroeconomics to microeconomics, or deny the importance of downward causation and the fallacy of composition. But by adopting the mainstream's misleading metaphor they are giving too much away; they are muddying these already murky waters.

Why it all matters

We need to remind ourselves again why all this is so important. It is not just a question of being careless with words. As noted at the beginning of the chapter, really important issues of economic policy are at stake. If all macroeconomic questions can be reduced to microeconomics, and there is no need to worry about the fallacy of composition, then austerity and wage cuts are easy to justify. The government, like the household, cannot live beyond its means, and must cut its spending to ensure that it does not. Mass unemployment means that the labour market has not cleared, and wages must be cut to make the quantity supplied equal to the quantity demanded.

All this is wrong, as earlier chapters have suggested. The government is not a household, and austerity often does much more harm than good. Across-the-board wage cuts reduce incomes and so lead to a decline in consumer demand. A falling price level would be very bad news. In macroeconomics, *ceteris* are not *paribus* (all other things are not equal). The attempt to reduce macroeconomics to microeconomics is a very bad idea, and when applied to policy decisions it can have very bad consequences.

Once again, this does not deny the relevance of microeconomics to macroeconomics (or for that matter vice versa); it is simply that the two bodies of knowledge exist side by side, neither being the foundation of the other. Macroeconomics is relatively autonomous, and certainly not independent

(compare Catalonia with Portugal, Wales with the Irish Republic, Puerto Rico with Canada). And there is mutual causation between macroeconomics and microeconomics, upwards and downwards.

There are some broader lessons for us all from the microfoundations delusion. First, we need to mind our language: we should think carefully about the metaphors we use, and make sure that they are helpful and not misleading. Second, we should pay serious attention to the methodology of economics: the philosophy of science is never going to be easy, but we neglect it at our peril. Third, we should be good neighbours: we should talk to political scientists, sociologists and anthropologists, and not talk down to them. They have dealt with the same issues of micro-reduction that confront us as economists, and they may well have something to teach us. If you have friends who are studying in these other departments, why not ask them what they think about these issues?

REFERENCES

Bedau, M. and P. Humphreys (eds) (2008), *Emergence: Contemporary Readings in Philosophy and Science*, Cambridge, MA: MIT Press.

Davis, J.B. (2003), *The Theory of the Individual in Economics: Identity and Value*, London, UK and New York, USA: Routledge.

Dawkins, R. (1976), *The Selfish Gene*, Oxford: Oxford University Press.

Dawkins, R. (1996), *The Blind Watchmaker*, Harmondsworth: Penguin.

Dawkins, R. (2006), *The God Delusion*, London: Bantam Press.

Dawkins, R. (2015), *Brief Candle in the Dark: My Life in Science*, New York: Ecco.

Gellner, E. (1956/2003), 'Explanation in history', in *Cause and Meaning in the Social Sciences, Selected Philosophical Themes, Volume 1*, London, UK and New York, USA: Routledge, pp. 1–17.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan.

Kincaid, H. (1998), 'Methodological individualism/atomism', in J.B. Davis, D.W Hands and U. Mäki (eds), *The Handbook of Economic Methodology*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 294–300.

Lakoff, G. (1996), *Moral Politics: What Conservatives Know that Liberals Don't*, Chicago, IL: Chicago University Press.

Lewontin, R.C., S. Rose and L.J. Kam (1984), *Not in Our Genes: Biology, Ideology, and Human Nature*, Harmondsworth, UK: Penguin.

Lukes, S. (1968), 'Methodological individualism reconsidered', *British Journal of Sociology*, **19** (2), 119–29.

Mandeville, B. (1714), *The Fable of the Bees or Private Vices, Public Benefits*, London: J. Roberts.

Mankiw, N. G. (2009), 'Back in demand: a great thinker has his admirers and detractors. Do his ideas logically cohere?', *Wall Street Journal*, 21 September.

Piketty, T. (2013), *Capital in the Twenty-First Century*, Cambridge, MA: Harvard University Press.

Piketty, T. (2019), *Capital et idéologie*, Paris: Editions du Seuil.

Ricardo, D. (1817), *On the Principles of Political Economy and Taxation*, London: John Murray.

Roosevelt, A.E. (1961), *The Autobiography of Eleanor Roosevelt*, New York: Harper & Brothers.

Skidelsky, R. (1967), *Politicians and the Slump: The Labour Government of 1929–31*, London: Macmillan.

Skidelsky, R. (1983), *John Maynard Keynes: Hopes Betrayed, 1883–1920, Volume I*, London: Macmillan.

Skidelsky, R. (1992), *John Maynard Keynes: The Economist as Savior, 1920–1937, Volume II*, London: Macmillan.

Skidelsky, R. (2000), *John Maynard Keynes: Fighting for Britain, 1937–1946, Volume III*, London: Macmillan.

Skidelsky, R. (2003), *John Maynard Keynes: 1883-1946: Economist, Philosopher, Statesman*, London: Macmillan.

Skidelsky, R. (2009), *Keynes: The Return of the Master*, London: A. Lane.

Skidelsky, R. and E. Skidelsky (2012), *How Much is Enough? The Love of Money and the Case for the Good Life*, London: A. Lane.

Smith, A. (1776), *The Wealth of Nations*, London: W. Strahan & T. Cadell.

Wickens, M. (2008), *Macroeconomic Theory: A Dynamic General Equilibrium Approach*, Princeton, NJ: Princeton University Press.

A PORTRAIT OF ROBERT SKIDELSKY (1939–)

Robert Skidelsky was born in China on 25 April 1939 to parents of Russian ancestry (part Jewish, part Christian). His father worked for the family firm, L.S. Skidelsky, which leased the Mulin coalmine from the Chinese government. He and his parents were interned in 1941, first in Manchuria and then in Japan; they returned to China in 1947, spending a year in Tientsin before moving to Hong Kong.

Between 1953 and 1958 Skidelsky was a boarder at Brighton College, and then studied history at Jesus College, Oxford (1959–61), UK. From 1961 to 1969 he was a research student and then Research Fellow at Nuffield College, UK. His first book, *Politicians and the Slump*, was based on his PhD dissertation (Skidelsky, 1967). Three years later he became Associate Professor of History at Johns Hopkins University, USA, but his early academic career was a chequered one. His 1975 biography of Oswald Mosley was felt by many to be too sympathetic with the British fascist leader, and he was refused tenure at Johns Hopkins, returning to the United Kingdom as Professor of History, Philosophy and European Studies at the much less prestigious Polytechnic of North London. In 1978 he was appointed Professor of International Studies at the University of Warwick and in 1990 he joined the Economics Department there, becoming Professor of Political Economy. He retains an emeritus chair at Warwick, and has been an Honorary Fellow of Jesus College, Oxford since 1997. His many honours include the life peerage that he was awarded in 1991, and appointment as a Fellow of the British Academy in 1994.

Skidelsky is best known for his magisterial three-volume, 1700-page biography of Keynes (Skidelsky, 1983, 1992, 2000); a single condensed volume (a mere 1000 pages) was published in 2003. The second volume won the Wolfson History Prize in 1992, and the third volume won another four prizes between 2000 and 2002. His short book, *Keynes: The Return of the Master*, provides an accessible introduction to Keynes's ideas and their application to the post-2007 Great Recession (Skidelsky, 2009). This was followed by *How Much is Enough? The Love of Money and the Case for the Good Life*, co-authored with his philosopher son Edward (Skidelsky and Skidelsky, 2012). Several more books have appeared since then, in addition to co-edited collective volumes on the question of who runs the economy, and the case for austerity versus that for a fiscal stimulus.

For an octogenarian, Skidelsky is still highly active. In late 2019 it was reported that he had recently written and filmed a series of lectures on the history and philosophy of economics, which was to be made available as an open online course in partnership with the Institute for New Economic Thinking, and was working on a book on automation and the labour market. He can still be hired as an after-dinner speaker, and among his declared interests (reported, as required by UK law, to Parliament) is a non-executive directorship of the Russian oil refining company Russneft.

Robert Skidelsky lives in Keynes's former country home at Tilton in East Sussex, and in 1991 chose Baron Skidelsky of Tilton as his title for the House of Lords. He went to the House of Lords as a member of the Social Democratic Party but soon joined the Conservatives and then, in 2001, became

➡

an unaffiliated cross-bencher. More recently he has shown considerable sympathy for the Labour Party under the leadership of Jeremy Corbyn. Skidelsky is a powerful critic of government austerity programmes in the UK and the European Union, and on 6 September 2019 he was a signatory with 81 others, including Victoria Chick and Thomas Piketty, of a letter published in the *Financial Times* defending Labour's proposals for increased government spending and a more egalitarian distribution of wealth through the establishment of an Inclusive Ownership Fund.

He has also made a convincing case against the micro-reduction project in mainstream macroeconomics, arguing that the teaching of macroeconomics should be taken out of the hands of microeconomists and incorporate ideas from politics, philosophy, and history. Readers should visit his website (www.robertskidelsky.com), which will supply them with thought-provoking material (every week or two) on current controversies in economics and economic policy.

EXAM QUESTIONS

True or false questions

1. The financial decisions of the government should not be made on the same principles as those applied by prudent individuals or families.
2. A falling price level is always good for aggregate output and employment.
3. In a closed economy, aggregate saving can only increase if aggregate investment expenditure also increases.
4. The use of metaphors in economics should always be avoided.
5. A cut in real wages is the only sure cure for mass unemployment.
6. In a closed economy, aggregate profits are determined by the spending decisions of capitalists on consumption and investment.
7. The methodology of economics is not something that an intelligent student should be concerned with.
8. 'Economics imperialism' has not resulted in the conquest of the other social sciences by mainstream microeconomics.
9. The same principles should apply to the finances of the nation and to the finances of each individual or family.
10. John Maynard Keynes took a strong interest in the methodology of economics.

Multiple choice questions

1. An increase in the real wage rate:
   a) never has any effect on aggregate output and employment;
   b) always leads to an increase in aggregate output and employment;
   c) has an effect on aggregate output and employment that depends on the reactions of capitalists and workers;
   d) always leads to a decrease in aggregate output and employment.
2. The use of metaphors in economics is:

a) entirely unproblematic;
b) so problematic that it should be avoided at all costs;
c) very uncommon;
d) common, interesting and useful, but also problematic.

3. The provision of 'macrofoundations' for microeconomics is:
a) a questionable use of metaphorical language;
b) entirely impossible;
c) absolutely essential;
d) unproblematic.

4. John Maynard Keynes:
a) took no interest in methodological issues;
b) endorsed the principle of methodological individualism;
c) would have rejected the case for microfoundations;
d) was an early supporter of neoliberalism.

5. 'Economics imperialism':
a) is a Marxist theory of the growth of colonial empires;
b) has succeeded in colonizing the other social sciences;
c) is advocated by heterodox macroeconomists;
d) has failed to colonize the other social sciences.

6. The 'paradox of liquidity':
a) is a meaningless expression;
b) is a fundamental principle of New Classical macroeconomics;
c) helps to explain the global financial crisis that erupted in 2007–08;
d) demonstrates the need to provide microfoundations for macroeconomics.

7. Scientific realism is:
a) a methodological principle accepted by most heterodox economists;
b) unattainable in the real world;
c) inconsistent with any formal theorizing in economics;
d) of no interest to economists.

8. Economists and practitioners of the other social sciences:
a) have almost nothing in common;
b) should stay well clear of each other;
c) should be encouraged to cooperate with each other;
d) agree on almost everything.

9. 'The fallacy of composition' is:
a) a problem encountered by landscape painters;
b) the source of some difficulty for economic theorists;
c) a mistake made by songwriters;
d) never a problem for economists.

10. The 'paradox of costs' is:
a) a problem in the neoclassical theory of the firm;
b) a difficulty experienced by accountants when studying company finances;
c) an argument for wage restraint;
d) an important proposition in Kaleckian macroeconomics.

Answers to the exam questions

Chapter 1

True or false questions

1) T
2) F
3) F
4) T
5) F
6) F
7) T
8) T
9) T
10) T

Multiple choice questions

1) c
2) b
3) d
4) d
5) a
6) a
7) d
8) b
9) c
10) c

Chapter 2

True or false questions

1) F
2) T
3) F
4) F
5) F
6) T
7) T
8) F
9) F
10) T

Multiple choice questions

1) b
2) c
3) b
4) c
5) d
6) b
7) a
8) d
9) d
10) d

Chapter 3

True or false questions

1) T
2) T
3) T
4) F
5) F
6) T
7) F
8) F
9) T
10) T

Multiple choice questions

1) a
2) b
3) c
4) b
5) b
6) b
7) b
8) b
9) c
10) b

Chapter 4

True or false questions

1) F
2) F
3) T
4) T
5) T
6) F
7) F
8) F
9) T
10) T

Multiple choice questions

1) b
2) c
3) b
4) d
5) b
6) c
7) b
8) d
9) b
10) c

Chapter 5

True or false questions

1) T
2) F
3) F
4) T
5) T
6) F
7) F
8) T
9) F
10) T

Multiple choice questions

1) d
2) a
3) c
4) b
5) b
6) c
7) d
8) d
9) c
10) b

Chapter 6

True or false questions

1) F
2) F
3) T
4) F
5) T
6) F
7) T
8) F
9) F
10) F

Multiple choice questions

1) a
2) b
3) d
4) a
5) c
6) a
7) a
8) a
9) d
10) c

Chapter 7

True or false questions

1) F
2) T
3) F
4) T
5) T
6) T
7) T
8) F
9) T
10) T

Multiple choice questions

1) c
2) b
3) a
4) b
5) a
6) d
7) c
8) a
9) c
10) b

Chapter 8

True or false questions

1) T
2) F
3) T
4) F
5) T
6) F
7) F
8) T
9) F
10) T

Multiple choice questions

1) b
2) c
3) d
4) a
5) b
6) c
7) d
8) c
9) a
10) d

Chapter 9

True or false questions

1) T
2) F
3) F
4) T
5) F
6) F
7) F
8) T
9) F
10) F

Multiple choice questions

1) a
2) c
3) a
4) c
5) b
6) d
7) d
8) a
9) c
10) c

Chapter 10

True or false questions

1) F
2) T
3) T
4) T
5) F
6) F
7) F
8) T
9) T
10) F

Multiple choice questions

1) b
2) d
3) b
4) c
5) b
6) d
7) c
8) d
9) a
10) d

Chapter 11

True or false questions

1) F
2) F
3) F
4) T
5) T
6) T
7) T
8) T
9) F
10) T

Multiple choice questions

1) c
2) d
3) c
4) a
5) a
6) b
7) d
8) b
9) a
10) d

Chapter 12

True or false questions

1) F
2) F
3) T
4) F
5) F
6) F
7) F
8) T
9) F
10) T

Multiple choice questions

1) c
2) d
3) b
4) b
5) b
6) b
7) b
8) b
9) b
10) d

Chapter 13

True or false questions

1) F
2) F
3) T
4) T
5) T
6) F
7) T
8) T
9) F
10) F

Multiple choice questions

1) c
2) a
3) b
4) d
5) a
6) c
7) b
8) c
9) a
10) b

Chapter 14

True or false questions

1) F
2) F
3) T
4) T
5) F
6) T
7) F
8) F
9) T
10) T

Multiple choice questions

1) d
2) c
3) c
4) c
5) b
6) c
7) a
8) b
9) d
10) a

Chapter 15

True or false questions

1) F
2) T
3) F
4) F
5) T
6) F
7) T
8) T
9) F
10) T

Multiple choice questions

1) c
2) d
3) c
4) a
5) a
6) b
7) d
8) b
9) a
10) d

Chapter 16

True or false questions

1) T
2) F
3) T
4) F
5) F
6) F
7) T
8) F
9) T
10) F

Multiple choice questions

1) c
2) a
3) d
4) b
5) a
6) b
7) b
8) d
9) a
10) c

Chapter 17

True or false questions

1) F
2) F
3) F
4) F
5) T
6) T
7) T
8) T
9) F
10) T

Multiple choice questions

1) c
2) d
3) c
4) a
5) a
6) b
7) d
8) b
9) a
10) d

Chapter 18

True or false questions

1) F
2) F
3) T
4) F
5) T
6) F
7) T
8) T
9) F
10) T

Multiple choice questions

1) c
2) b
3) a
4) d
5) b
6) a
7) d
8) d
9) c
10) a

Chapter 19

True or false questions

1) F
2) T
3) F
4) F
5) T
6) T
7) T
8) T
9) T
10) F

Multiple choice questions

1) b
2) c
3) c
4) a
5) c
6) a
7) a
8) d
9) b
10) a

Chapter 20

True or false questions

1) F
2) F
3) T
4) F
5) F
6) T
7) F
8) T
9) T
10) F

Multiple choice questions

1) d
2) d
3) d
4) a
5) b
6) d
7) b
8) c
9) c
10) d

Chapter 21

True or false questions

1) T
2) F
3) T
4) F
5) F
6) T
7) F
8) T
9) F
10) T

Multiple choice questions

1) c
2) d
3) a
4) c
5) d
6) c
7) a
8) c
9) b
10) d

Index

absolute competitive advantage 438, 448–50
absolute income hypothesis 242, 245, 257–8
ABSs *see* asset-backed securities (ABSs)
accelerator effect 388
accelerator models 268, 285, 298
accelerator principle 283–5
accounting decomposition of profit rate 299
Accumulation of Capital 466
adding-up constraint 461
adding-up theorem 84
adjustment costs 272–3
adjustment mechanisms 487, 541, 552–3, 559
adjustments
  crisis-induced 541, 553, 557, 558
  external 553–60
  price and quantity 88
  symmetric 557–9
advanced economies 379–82, 395
The Affluent Society 576
Aftalion, A. 283
agents *see* economic agents
aggregate consumption 245–7, 255, 258–9, 596
aggregate demand 307–27
  and aggregate supply, Keynes on 423–4, 426
  bank lending and expectations of 135–6
  consumption and 234, 241, 280–83
  definitions 307, 308–9
  demand management policies 325–7
  expected 323–5
  firms and production plans 131–2, 140
  fiscal policy and 28, 194, 356
  government intervention 34, 52
  income multiplier 321–3
  Kaleckian model 373, 530
  Keynesian economics 311
    economic growth and 376, 384–6, 387
    macroeconomics approach 387, 411–13, 426
    output and 234, 315–21
    policy guidance 427–8
  monetarists' approach 411–13
  neoclassical theory 312–14
  subcomponents 309
  subdivision of 316
aggregate liquidity 607
aggregate profits 424, 607
aggregate supply
  and aggregate demand, Keynes on 423–4, 426
  expected 323–5
  neoclassical model 314–15
agriculture 77, 93–8, 382
alienation 349–50
allocation of resources 87, 447, 453, 518
'alternating movements of creation and cancellation of money' 137
amortization 318–19, 350
animal spirits 33, 42, 70, 240, 277, 389–90
anti-growth bias 222, 500–501
Appalled Economists 31
Appelbaum, E. 531, 532
applied theory 105, 106–7, 109, 112, 115, 118
Argentina 183, 442, 448–9, 454, 553
Asian crisis (1997–98) 477, 478, 553–4
Asian economies 379–81
asset-backed securities (ABSs) 525
assets
  bank balance sheets 162–8, 216
  hedging 178
  real 318
  savings and 186
  securitization and 529–30
  wealth and 403
assumptions
  of comparative advantage theory 447
  of demand theory 256–60
  in modelling 38–40
  theory of monetary circuit 130–31
asymmetric information 183–4, 273
austerity policies 2, 27, 29, 51, 52, 57, 387, 451, 481, 557, 595
  see also fiscal consolidation
Australia 445, 448–9, 457, 603
Austrian school 536, 550

autarky 441, 442, 467
automatic stabilizers 355, 363–5, 558
autonomous consumption 235, 242, 280–81, 316–17
autonomous demand 387, 390, 392, 396
autonomous factors, affecting bank reserves 167, 169
average propensity to consume (APC) 242–6, 249, 251–2, 257–60
average wage rate 90
Ayres, C.E. 572, 575, 582

bad economic models 38
bad economics 26
balance-of-payments
  accounting identity 473
  adjustment 466
  crisis 474, 495, 505
  deficits 484, 554, 555, 559
  equilibrium 14, 469, 472–3, 474–5, 480, 482–5
  problems, Thirlwall's interest in 490
balance-of-payments constrained growth rate 469–88
  definition 470
  determining 473–80
    diagrammatical approach 481–6
    including growth of capital flows 480
  heterodox view 470, 472–88
  mainstream view 470–71
  non-price competitiveness in international trade 469, 470, 479–80, 482–3, 485, 487–8
  price competitiveness in international trade 469, 475–7, 479, 488
  tests of model and empirical evidence 487–8
  Thirlwall's Law 469, 470, 477–9, 484, 487, 488
balance sheets 163–6, 168, 170, 176, 185–7, 216, 345–6
balanced budgets 366, 371
balanced trade 447, 466
banana diagram 291–3
bancor 495, 496, 497
Bank Act (1935) 549
Bank for International Settlements (BIS) 169, 213, 507, 523
bank notes 155, 157–8, 204
Bank of England 157, 211
Banking School 118, 172
bankruptcy 212, 503, 523, 552
banks/banking 151–70
  asymmetric information 183–4
  bank credit and money creation 132–4, 142, 162, 170, 549
  decision to lend 134–6
  deposits 153–6, 158–60, 162–3, 178, 197, 203, 215, 276, 343
  failures 158, 164, 179, 180, 183
  financial crisis and pessimism of 136, 143
  flux/reflux principle 152, 159–61
  golden rule 344–6
  as heart of economy 92
  heterodox perspective 156–61, 169–70
    and modern payment system 161–9
  importance in macroeconomic analysis 152
  institutions 153–4, 155, 157
  loans 131, 137, 139, 169, 178–9, 215 *see also* credit
  mainstream view 151, 152–6
  means of payment 152, 156–7, 160, 170, 504, 512
  monetary circuit and 133, 134, 143, 152, 159, 172
  near collapse of international 183
  payments and settlements systems 152, 163–5
  uncertainty 134–6, 212
  see also central banks; commercial banks/banking
barter
  economy model 129
  as 'imagined state' 126
  reasons to 127
  relationship with money 123–4, 126, 128, 153, 157, 170
barter theory of money 122
barter vision of trade 448
Basel Agreements 213
Batt, R. 531, 532
beggar-my-neighbour policies 453–4, 466
behavioural finance 184–5
Belgium 404, 497, 512
Bernanke, B. 38, 61, 65, 184, 274
Big Bank 553
Big Government 553
black swans 254
Blanchard, O.J. 54–6, 58, 64–6, 68, 70–71
Blecker, R.A. 461–2
The Blind Watchmaker 602
bonds
  in 2008 financial crisis 180–81
  ABSs, CDOs and CDSs 525–6
  central bank independence and 500
  China and US 448
  definition 176
  ECB 498, 499
  government and budget deficits 362, 368–9
  historical view 189–90
  in Keynes's analysis 193, 194

liquidity 276
in modern finance 194–5
in quantitative easing 201, 211, 326
secondary markets 178–9
as technique for managing liquidity 183
book-entry nature of money 215–16
Borio, C. 214, 224
borrowers
asymmetric information 273
creditworthy 133–6, 155, 162, 170
as deficit-spending units 549
'fringe of unsatisfied' 135
households 183
net worth of 184
non-performing loans 169
preference
for extended periods of time 187, 188
liquidity 212, 275–6
pull-backs 553
risks 293, 552
in 'speculative finance' position 551
Bortis, H. 84, 97, 98, 102–3, 107, 111, 113, 114, 118
Boulding, K.E. 567, 570, 577
bounded rationality 254
bowl and marble analogy 32, 33
Brazil 27, 183, 377–8, 444, 445, 454, 553
Bretton Woods 15, 146, 203, 303, 497, 506–7, 512
Brexit 456–7
broadly based standard of living 566, 574, 575
Brundtland Commission (UN) 571
bubbles
accelerator and 523
banks and 212, 220, 222
credit 210
definition 200
dot.com 180, 210, 288, 300
real-estate 61, 210, 212, 220, 478, 503, 525, 529, 553
in wealth channel 219–20
budget deficits 355, 360–63, 366–72, 481, 546
budget surplus 363, 366
building metaphor 611
Bundesbank 206–7
'bursts of optimism and pessimism' 34, 122, 143, 318
business cycles 18, 91, 105, 111–13, 192, 194, 308, 548, 560

Caballero, R.J. 114, 222
Cambridge capital controversy 77, 103, 275, 398
Canada 166, 168, 258, 445, 446, 448–9, 454
Candle in the Dark 603
capacity effect 112, 268, 289–90, 303
capital accumulation
distinction between net and gross investment 269
explanations for slowdown 297, 300
factors causing difficulties in 427
growth of demand for goods and services and 472
impact of financialization on 530–31
as increase in stock of investment equipment 270
Keynes's optimistic view about economic growth by 577
monetary economy of production based on 341
rate of 290, 292–3, 472, 481
reduction in fixed 350
Robinson's model 466
utilization and profit rates 299
world aggregate demand and supply encouraging 428
Capital and Ideology 605
capital flows 421, 456, 473, 474, 478, 480
capital formation
logic of fixed 347–9
process of 346, 349–50
'capital imports' 554
Capital in the Twenty-First Century 605
capital stocks 566, 568, 574, 579–82, 585, 590
capital-theoretic debate 104, 113
capitalism
financialization and 520–21
financing needs of modern 188–90
long record of instability 541
Marx and 430
monetary disorders and 347
underconsumption tendence 543
capitalist economies
excess supplies and 540
full employment and 371
indebted, Minsky on 197
inherently cyclical growth pattern 544, 548, 550
instability of 179, 194, 550
Kalecki's perspective 373
large financial sectors in 177
long-term debt markets affecting structure of 191
middle class membership 192–3
paramount characteristic of 137
Polanyi's perspective 536–7
see also advanced economies; monetary economy of production
capitalist reproduction 192

capitalists 40–41, 188–90, 191–3, 292, 296–7, 543–4, 609
Capitalizing on Crisis: The Political Origins of the Rise of Finance 519
carbon tax 571, 573
cartels 91–2, 191, 520
catching up 380–81
cause-effect relations 594
CDOs *see* collateralized debt obligations (CDOs)
CDSs *see* credit default swaps (CDSs)
Cencini, A. 343
central banks 200–225
  Austrian school 550
  budget deficits and 369
  EPU and 507
  essential role of 200
  exchange rate regimes and 209–10, 556
  independence 221–2, 499–500, 502
  interest rates and 59, 206–9, 211, 218–20, 223–5, 274, 319, 359, 360
  intervention in aftermath of global financial crisis 334
  as lenders of last resort 553
  monetary policy 274, 319, 335, 557
    credibility 221–2
    demand management policies 307–8
    ECB 499–500, 502
    heterodox perspective 213–24
    instruments 326
    mainstream perspective 202–13, 224–5, 335
    from post-Keynesian perspective 214–24
  reserves 202–3, 204, 211, 216
  TARGET2 system and 504, 506, 508
  Taylor rule 59, 201, 208, 223, 224
  see also Bank of England; European Central Bank (ECB); Federal Reserve
ceteris paribus condition 35
change 35–6
  see also institutional change; structural change
cheques 163, 188, 204
Chicago School 399
China 377, 416, 439, 443–4, 447–50, 453, 455–9, 461, 481, 521, 558
choice
  free 340–41
  portfolio 181
  post-Keynesian theory of consumer 578
  of production factors 284
  between work and leisure 405, 412
  see also intertemporal choice
circular flow
  central bank money 508–9
  definition 407
  illustration 406
  of income 390, 610
circular process of production *see* social and circular process of production
circulating capital 99, 158, 188–9
Clark, J.B. 408–10
class-based approach 130–31
classical economics
  and law governing factor payment 420–22
  proponents of 377, 407, 430
classical–Keynesian political economy 97, 102–3
  assessment 113–14
  capital-theoretic debate 104
  counter-revolution 100–102
  cycles and trend 111–13
  determination of long-period or trend output and employment 109–10
  formation of prices 107–8
  principles and theory 105–7
  supermultiplier 110–11
  system outline 102–13
classical labour theory of value 77, 403, 422
'Classics' versus Keynes 189
clearing balances 165, 166
closed economies 244, 357, 387–9, 607–8
'cobweb theorem' 399
cognitive dissonance 42–3
collateral 211, 217
collateralized debt obligations (CDOs) 180, 525
Collective Choice and Social Welfare 589
colonialism 438–9
commercial banks/banking
  balance sheet example 216
  central banks and interest rates 274
  definition 152
  history 157, 170
  inflation example 345–6
  monetary base and money multiplier 202–3
  public sector transactions examples 167–9
  reserves at central bank assumption 161–2
  in triangular nature of money illustration 217
commodity money 153–5, 156
common currency 496
common good 30, 225, 578

common stock *see* shares
companies *see* firms/companies
comparative advantage
in countries with abundant natural resources 460
definition 438
dynamic 451
example 440–42
traditional theory of 440–47, 448
static nature of 451, 466
competition
classical economics 100, 101, 421, 431
coercive 296
financial investment reducing 278
free 520
imperfect 76, 92, 466
in manufacturing sector 95–6
over global trade 445, 452, 462–3
Sraffa on 76
see also perfect competition
competitive advantage *see* absolute competitive advantage
competitive markets 90–91, 93, 544–5, 596
The Complete Works and Correspondence of David Ricardo 76
Comprehensive and Progressive Agreement for TPP (CPTPP or TPP-11) 457, 459
conflict 41
conflict theory of inflation 335
conservative bias 31
consumer choice 578
consumer demand 256, 263, 342–3, 612
consumer price index (CPI) 339–40
consumers 235–6, 245–6, 250–51, 253–60, 263, 333–4, 445
consumption 233–61
consumption functions
empirical testing of 244–6
Keynesian 234, 241–4, 596
life-cycle hypothesis 234, 247–50, 260–61
mainstream 235–8
permanent income hypothesis 250–52
policy implications 260–61
random walk 238
short-run and long-run 245–6, 252, 255, 257–9
current monetary policy and 213–14
Duesenberry and 263–5
expenditures 233, 238, 241–4, 248–9, 256–8, 544
fiscal policy and 234, 235, 243, 261
fixed capital formation and 347–9
heterodox perspective 253–60, 261
income and 234–5, 236, 238, 240–52, 255–61
inflation and 334, 340
investment and 280–83, 287, 292, 300
Keynes's approach to 238–44, 246, 247, 250, 256
mainstream theories 247–53, 259–60, 261
in monetary circuit with state 142
present value of lifetime 236–8, 248
random walk theory of 252–3
spending power and 81, 111
taxation and 244
see also aggregate consumption; autonomous consumption; luxury consumption; marginal propensity to consume (MPC); private consumption; public consumption; underconsumption
contagion 478, 553–4
convergence criteria (Maastricht) 498–9
convex adjustment costs 272
corn model (Ricardo) 76–7, 81, 99
coronavirus *see* Covid-19 pandemic
corruption 528
cost minimization 87
cost of living 339–40
cost-push inflation 335, 547
cost(s)
adjustment 272–3
debt servicing 551–2, 558
of development 582
impact of technological progress on 338–9
of living, inflation and 339–40
paradox of 607
of production 76, 93, 101, 133, 138, 192, 227, 314, 321, 345–6, 407, 430, 450, 461
see also external costs; financing costs; labour costs; marginal costs; opportunity costs
counterparty deposits 158
Covid-19 pandemic 29, 52–3, 459
credit
banks' decision to lend 134–6
central role in financing of production 158–9
consumer 528
creation, limits to 169
economic growth and 548–9
inflation and 222, 334–5
interest rates and 210, 274, 326
money creation and 132–4
overextension 550

as policy instrument 326
pro-cyclical behaviour 548
credit bubbles 222, 502–3, 523
credit cards 524, 528
credit crunch 113, 552–3
credit cycles 541, 548–9, 550, 551, 560
credit default swaps (CDSs) 526
credit derivatives 523
credit flows 503
credit market imperfections 273
credit-money
creation 158, 160–61, 170
endogenous 549
credit rating agencies 503, 505
credit standards 211
creditor countries (euro area) 502–3, 505, 508
creditworthiness 133–4, 136, 162, 164, 295
crisis-induced adjustments 541, 553, 557, 558
crowding in 355, 361
crowding out 315, 355, 361, 368
culture 536, 579
cumulative causation 270, 399, 490
currency 157–8, 204
appreciation 444
common versus single 496
crisis 474, 554
depreciation 326, 555
devaluations/revaluations 559
elastic 549
exchange rate regimes and 556
purchasing power and 335–6, 339–40
single (EU) 495–6, 498, 502, 508
current account 473–80, 554, 556–8
deficits 472, 473–4, 479, 554–5, 558
surpluses 321, 449, 481, 503, 558
current expenditures 356, 357
customs unions 454, 457
cyclical growth dynamics 541, 547–53, 563
cyclical growth pattern 540, 547–8, 551
cyclical movements 91, 111–13
see also business cycles
cyclically adjusted budget 366

Dallery, T. 299
Daly, H.E. 571, 574, 581, 584
Das Kapital 99–100, 430
Davidson, P. 20, 38, 172, 193, 254, 460, 561, 577–8
Davies, R. 527, 531
Dawkins, R. 602, 603, 604, 607–8
debt
business cycles and 197, 551
crises 553–4
definition 177
financial instability and 549
financialization and 177, 502, 522–3, 528, 530
formation, production as process of 133
impact of saving 138, 141
long-term 191, 193
money as spontaneous acknowledgement of 343
from national income 177
on purchase of goods 216
reimbursement 135, 138, 141, 142, 346
short-term 160–61, 180–81, 195, 551
see also external debt; indebtedness; net debt; public debt; sovereign debt crisis
debt deflation spiral 553
debt-led booms 450, 502
debt service payments 197
debt servicing costs 551–2, 558
debt structures 195, 197
debtor countries (euro area) 508, 509
debtors (Ponzi) 551–2
decision making 186, 427, 499–500, 578, 579, 596
deferred consumption 313
deflation 212, 337, 342–3, 347, 351, 481, 523, 595
deflationary shocks 559
deindustrialization 377, 382, 444
Delors Committee 497–8
demand
increasing 334
investment as sensitive to changes in 140
sources of, in long run 386–7
see also aggregate demand; consumer demand; effective demand; export demand; supply and demand; total demand
demand curve 87–8, 92, 104, 187, 235, 479
demand effect 268, 289–90, 299, 303
demand-led growth 376, 384–6, 396
demand management policies 307–8, 315, 325–7
demand price 88
depressions 32, 180, 548, 558
see also Great Depression
deregulation 180, 456, 519, 599
derivatives 524, 528
Diamond, D.W. 183
direct labour 81, 97, 98, 99, 107–8, 422
'discipline in the factories' 609
disembodied firms 349

disequilibrium 56, 91, 289–90, 396, 542
disposable income
consumption and 234, 241–3, 247–8, 257–9, 316–17
definition 234
in elaborated multiplier 322, 323
household savings and 318, 360
paradox of thrift and 318
in simple income multiplier 321
social benefit and 320
taxes reducing 320
distribution
international trade and 445–7
Sen's interest in issues of 589–90
surplus principle of 81, 98–9, 103, 118
see also income distribution; income redistribution
distributive justice 83, 99, 103
divergence
in euro area 503
income 392, 395, 396
in national productivity and output growth rates 471
between rich and poor countries 376, 381
division of labour 153, 173–4, 393
Doha round 455
dollar (US) 336, 449, 497, 512, 556, 559
dot.com bubble 180, 210, 288, 300
'double coincidence of wants' 125–6, 127, 153
double-entry bookkeeping 157, 158, 337–8, 343, 348
double movement 537
'downsize and distribute' policy 524
downward causation 594, 597, 604–8, 610, 612
Drakopolous, S.A. 239, 240, 256
DSGE (dynamic stochastic general equilibrium) models 56
argument over 55–8
ECB casting doubt on 72
lack of alternative to 57, 70
mainstream view 58, 60–61
rejecting 66–9
rethinking 62–6
Duesenberry, J.S. 246, 255–60, 263–5
Dutch disease 444
Dybvig, P.H. 183

East Asia 440–42, 451–2, 460
ECB *see* European Central Bank (ECB)
Ecofin Council 499
ecological economics 576–9
econometric models/studies 241, 246, 476, 530–33, 598
economic activity
aggregate demand as driving force of 133
automatic stabilizers and 363–5
budget deficits 360–61, 362, 367, 368
finance as most globalized aspect of 554
fiscal policy affecting 356
as governed by supply factors 101
influenced by external shocks 91
investment and 281, 288, 361, 373
principle of effective demand 81
scale of, and employment 96, 211
as supply-determined 85
economic agents
behaviour 34, 84, 91, 339, 347 *see also* rational behaviour; individual behaviour
inflation and 339
rational expectations 221, 252, 255, 261, 600–601
structural monetary reforms and 347
unemployment and 339, 342
see also capitalists; consumers; firms/companies; government(s); household(s); intermediaries/intermediation; rentiers; workers
economic consequences of long-term finance 190–93
The Economic Consequences of the Peace 145
economic crises
countertendencies used by entrepreneurs 543–4
Covid-19 52–3
monetary policy-making 222
need for heterodox approach to analysis of 3
possible sources of 141, 143
robustness of welfare state 323
role of state in monetary circuit 141–2
see also financial crises
economic development
versus economic growth 573–5
economic growth equated with 377–8, 567
heterodox economics and true 575–9
neoclassical model of 568–9
Sen's research into 589–90
economic growth 382
aggregate demand and 376, 384–6, 387
central bank policies 274
credit and 548–9
cyclical 548, 568
definition 376, 377
demand-led 376, 384–6, 396

versus economic development 573–5
economic development equated with 377–8, 567
European 500, 501, 502, 509
export-led 439, 446–7, 461, 482, 501, 502
financialization and 533
heterodox view 472–88
income distribution and 41, 377, 392–3, 398, 466, 546–7, 560, 596
international trade and 438–9, 460
Keynesian growth theory 383–4, 387–96
linked to theory of money 128
mainstream view 382–4, 470–71
natural rate of 205
neoclassical model of 568–9, 589
as public policy issue 567
quantitative easing 201, 211, 212
savings, investments and 85, 86, 290
secular stagnation and 29
Solow model 470–71, 568–9
supply-led 377, 382–4, 396–7
sustainable 223
sustainable development and 567–77, 579–85
trickle-down effect 574
unsteady 377, 380, 382
see also balance-of-payments constrained growth rate; long-run economic growth
economic illness 96
economic integration 454, 456, 512
economic policies
austerity as 27, 51, 52
effectiveness related to consumption function 234, 260–61
fiscal policy as 355, 373
implications of microfoundations acceptance 595
importance 26
political science and 42
pure theory representing basis for 107, 115
relevance 36
see also fiscal policy; monetary policy
'Economic possibilities for our grandchildren' 567
economic power 427, 519, 572
economic rationality 235, 253–6
economic theories
further reading 114–15
history of 82–115, 396
importance 82
laws of 37
neoclassical–Walrasian and classical-Keynesian assessment 113–14
political economy 92–113
two broad groups 82–4
economics
as about markets governed by immutable natural laws 26–7
approaches to 25–31, 43–5
'business as usual' 53
crisis of 3
definitions 3, 44–5
depression 54, 58
history of 84–92, 596–7
ideology in 31–4
need for heterodox approach 1–3
philosophical foundations of 609
pre-World War II 422
saltwater versus freshwater 55
as science argument 34–7, 60
social sciences and 26, 40–43, 605, 611
studying
purpose of 1
stages of 1–2
use of mathematics and models 37–40
wealth sharing at heart of 403
see also classical economics; heterodox economics; Keynesian economics; mainstream economics; neoclassical economics; post-Keynesian economics
economics approach 82–3, 84
economics imperialism 604, 605
Economics of Development 491
The Economics of the Short Period 328
economies of scale 322, 460
The Economist 55, 60, 62, 63
economist approach to monetary union 496–7
economists *see* master economists; *individual economists*
effective demand 423
as barrier in labour market equilibrium 90
cycles and trend 111–13
definition 308
expected aggregate demand and supply 323–5
growth rate of 110
Keynes's view of 91, 423–4
level of employment governed by 83–4, 89–90, 324, 423, 595
principle of 81, 84, 90, 101–5, 109, 373, 398, 596, 598
supermultiplier and 109–10
efficiency
in allocation of resources 447, 453, 518

of capital, marginal 276, 403, 410, 425, 426
sustainable development and 572
in trade agreements 454–5
efficiency wage 547
'efficient allocation of scarce resources' 3, 44–5, 405, 518
Egypt 378
Eichner, A.S. 30, 227, 294, 296, 560, 561
elaborated multiplier 322–3
'emancipation of the mind' 82, 115
embeddedness 43, 536, 537
emergent properties 594, 606–7, 610
employment
aggregate demand and 234, 309
changing composition of 382
effective demand and 83–4, 89, 324, 423, 595
factors of production and 91
financialization and 531–2
income distribution and 83
income from 248, 250
investment and 190–94, 278, 297, 318, 424–6, 533
levels 211, 219, 221–2, 238, 315, 371, 546, 595
macroeconomic theories 311, 312–13, 314–16
maximization 223
monetary economy of production and 341, 423
offshoring, trade and wage inequality 446–7
output and 90–91, 96, 101–2, 105, 107, 109–13, 114, 190–93, 221–4, 315, 594, 597
as political issue 96
stabilization 224
technical change and 428
young people's prospects of 528
see also full employment; unemployment
empty money 344–5, 346, 350, 351
'The end of the business cycle?' 53–4
endogenous credit-money 549
endogenous cycles 297
endogenous growth theory 471, 569
endogenous money creation 118, 549
Engels, F. 430
entrepreneurs 85, 88, 90, 111, 270, 277, 300, 419–21, 423–9, 544, 571
environment
economy and society, as interrelated 577, 580–82, 585
impact of economic growth on 567, 570, 576, 579, 585
organism relations and 604
sustainability and 569–75
environmental damage 444, 459, 578
environmental value 583
Epstein, G.A. 520, 521, 522, 523
EPU *see* European Payments Union (EPU)
equality 567, 579
see also inequality
equation of exchange 204–6
equilibrium *see* macroeconomic equilibrium; market equilibrium
equilibrium analysis 69, 89, 336, 399, 579
equilibrium price 87–8, 215, 396
ergodic systems 578
Essay on Profits 77
An Essay on the Influence of a Low Price of Corn 118
Essays in Persuasion 577
Essays in the Theory of Economic Growth 466
Essays in the Theory of Employment 466
Euler's theorem 410
euro area
budget deficits 367
central bank interest rate 360
currency 496
ECB's governance 498, 499, 500–502
financialization and 502–3
increasing competitive advantages 449
internal devaluation 450
Maastricht criteria 495, 498–9
Parguez's critique of 172
payments 505–7
rebalancing mechanism 508
recommendations for 508–9
TARGET2 system 504–6
see also financial crises: euro area
Europe
banks lacking trust 164
demand formation 386–7
economic stagnation hanging over 181
passing Companies Acts 189
unemployment periods 325
see also individual countries
European Central Bank (ECB)
additional reserves created by 167
doubts about DSGE model 72
governance of euro area 498, 499, 500–502
monetary policy stance 222, 499–501, 502, 508–9
relation to TARGET2 system 504, 507
as supranational central bank 172
European Currency Unit (ECU) 497
European Monetary Cooperation Fund 497

European Monetary Institute 498
European Monetary System (EMS) 497, 553–4, 559
European Monetary Union (EMU)
heterodox perspective 502–9
hostile environment for government finance 180
mainstream perspective 495–502
Stability and Growth Pact 366, 502
European Payments Union (EPU) 495, 507, 512
European Union (EU)
budget positions of member states 366–7
ECB and 500
inflation rates 498
new economic nationalism 456–7
trade agreements 454, 455
see also individual countries
Eurostat 499
'euthanasia of the rentiers' 194, 426
'evolutionary stable strategies' 603
ex nihilo money creation 158–9, 162, 169, 172
excess supplies 336, 540, 543, 545, 556, 595
exchange economy 83, 86, 129–30, 134, 182
exchange rate(s)
in Asian financial crisis 478
devaluation/depreciation 209, 476–7
fixed 495–7, 556
flexible 326, 556
impact on balance of payments 555
overvalued 382
pass-through 210
regimes 556
stability 498
see also real effective exchange rate (REER)
exogenous givens 399
exogenous shocks 32, 58, 542
expansionary fiscal consolidation 362
expansionary fiscal contraction 29
expansionary policies
abandonment of 28–9
demand management 315, 325
fiscal 34, 43, 315, 326
macroeconomic 460, 484
monetary 211, 212, 274, 315
see also quantitative easing (QE)
expected utility 239, 251, 253
expenditure flow 159, 160
export demand 376, 390–91, 395, 399, 487
export demand function 470, 476
export income elasticity of demand 470, 475–88
export-led growth 439, 446–7, 461, 482, 501, 502
export price elasticity of demand 477
exports
balance-of-payments constrained growth 481–6
international trade 320–21, 326, 442–5, 447–8
in Kaldorian growth theory 390–91
manufactured 460–62
net 309–10, 411, 546, 555
trend 109–10
see also international trade
expropriation 352, 527
external adjustment 540–41, 553–60
external costs 570–71, 583
external debt 352, 554, 558
external deficits 554, 555, 557, 558, 559–60
external shocks 91
extra-market agents 546

Fable of the Bees 607
factor income 308–10, 313, 316–18, 321, 322, 323
factor payments
law governing 420–22
product 'exhaustion' and 409
factor substitutability 271
factors of production 84, 87, 89, 91, 104, 310, 312, 382–3, 404, 408, 419–21, 445, 470–71
failure hypothesis 382
fallacy of composition 17, 20, 461, 545, 593–4, 606–7
falling behind 380–81, 392, 395, 396
fascism 536–7
Federal Reserve 62, 180, 210–11, 214, 220, 501, 550
feminist economics 579
fictitious commodities 536–7
final finance 160–61
final payment 201, 215, 216, 504–8
'finance capital' 191, 520
Finance Capital 520
financial capital 224, 503, 554
financial crises
2007–08 180–81
examples of 'fundamental uncertainty' 253
expansionary fiscal policies 43, 486
failure of economists to predict 2, 38–9, 57, 61, 64–6, 71
financial sector trends following 522
financialization and 518, 527–8
inability of economists to propose policies to deal with 52, 57
inflation and 334
mainstream economics, civil war 51–2, 55, 62

Mankiw's theories and policies 31, 55
near zero interest rates since 195, 219
neoclassical theory and 123
'network theory' 185
as paradox of liquidity example 607
as possible Minsky crisis 552
quantitative easing as misguided policy 151, 172, 211
relationship between financial markets and economic growth 533
situation prior to 53–4, 58
TINA's flawed logic 27–8
unravelling of demand-generating process 386
world trade in goods and services levelling off after 439
American/subprime 180–81, 558
characterized by failure of private sector debt 179
early signs 62
financialization and households 527–8
inflation-targeting strategies 220
low investment since 300–301
mainstream explanation 210
mortgage-backed securities 525
overproduction as trigger for 547, 552
as well-known black swan event 254
capitalism's propensity for 544
dynamics, international 553–4
era of 'Great Stagflation' 598
euro area 180–81, 495
creation of additional reserves 167
deficit countries 450–51
ECB policy stance 501–2
government debt crisis 179
inflation-targeting strategies 222
internal devaluation 450
PIIGS countries most affected by 503
recommendations for 508
roots of 502
'secular stagnation' and 29
TARGET2 balances 504–5
Minsky's consideration of 197, 294
outbreak since 1980s 288
paradox of tranquility 294
as rebalancing adjustment mechanism 552–3
systemic 202, 213, 502
see also Asian crisis (1997–98); balance-of-payments: crisis; crisis-induced adjustments; depressions; Great Depression; Great Recession; recessions; sovereign debt crisis
financial deepening 160–61
financial economics
mathematical approach 188
textbook approaches 187, 191, 193
financial engineering 529, 532
'financial expropriation' 527
financial fragility 51, 58, 212, 551–2
financial innovation 190, 523, 524, 552
financial instability 184, 194, 223, 502, 530, 541, 548, 549–50
financial instability hypothesis 70, 197, 551–2
financial institutions
balance sheets 185
financialization and 502–3, 527
payments systems 164
profits 522
'too big to fail' 212
see also banks/banking
financial intermediation 160, 187–8, 346, 549
financial markets
differences in interest rates 188
financial deepening and 160–61
financialization and 502, 518, 521, 528, 531, 533
international integration 179
productive investment and 279
short-termism and 527
as subject to bouts of irrational optimism and pessimism 277
Tobin's Q and 278–9
financial regulation 161–2, 534
financial savings 138–9, 142, 313–14, 317
financial stability 197, 202, 210, 219, 223–4, 502
financial system 176–96
finance
'Classics' versus Keynes 189
consequences of long-term 190–93
modern 194–5
textbook approach to 185–8, 191
financial innovation and 524
heterodox perspective 185–90
Keynes's analysis 193–4, 424
mainstream perspective 181–5, 189
size and activities of 177–9
financial wealth 177
financialization 517–34
definitions 502, 518, 519–21
dimensions 521–3

economic growth and 532–3
employment, human capital, research and development, wages and 531–2
euro area 502–3
households and 527–8
impacts 528, 530–33
income distribution and 532–3
investment and 299–301, 530–31
macroeconomic models and 529, 530
neoliberalism and 518, 521, 533, 600
non-financial corporations and 522, 523–4, 526–7, 531
periodization 521
policies/programmes reducing 534
financing costs 190, 192, 273, 274
financing needs of modern capitalism 188–90
financing structures 182, 197
fine tuning (fiscal) 364
firms/companies
banks and 211, 214, 218, 219, 221, 223
behaviour 33–4, 42
champions 452–3
Companies Acts 189
disembodied 349
in financial system 177, 186–7, 188, 190, 191–3, 195
financing 189–90
government spending 142–3
household savings and 138–9, 141, 193
investment 269–70
empirical testing 297–300
heterodox perspective 289–94, 296–7
Keynes's approach 276–85, 287, 288
mainstream view 271–5
reluctance 300–301
in monetary circuit 131–43
post-Keynesian theory of 294–6, 298
private 131, 309, 315, 317–18, 326
see also multinational corporations (MNCs); non-financial corporations
'fiscal compact' 366
fiscal consolidation 3, 59, 362, 495, 501, 508–9
see also austerity policies
fiscal contraction, expansionary 29
fiscal equalization mechanism 503, 504
fiscal policy 355–71
abandonment of 59
aggregate demand and 28, 194, 356
consumption and 234, 235, 243, 261
countercyclical 145
critiques 368–71
definition 355
expansionary 34, 43.315, 326
functional finance and 365–8, 370
government expenditure and 356, 357, 361–3, 364, 366, 368–71
lags in application of 364
mainstream view 358–9, 361–2, 371
need for 356–7
post-Keynesian perspective 360–62, 365, 371
role of automatic stabilizers 363–5
undertaken by governments 326
fiscal profligacy 367, 501
Fisher, I. 235, 238, 241, 247, 252, 273, 553
fixed capital 99–100
formation 347–50, 352
goods 100, 347
investment 158
flexible exchange rate 326, 556
flexible price system 182
flux 152, 159–61
forecasting 364, 598
foreign direct investment (FDI) 474, 480
foreign trade 316, 544, 561
forging ahead 380–81, 392, 395, 396
fractional reserve banking 154, 161
France 31, 93–4, 129, 269, 367, 522
Franco-Italian Circuit School 172
fraud 528, 573
free market economics 34, 600
free market economy 265, 430
free market(s) 430, 537
free trade 441, 445, 451–2, 454, 459, 462, 466–7
French Keynesian economists 172
freshwater economics 55
frictional unemployment 332, 340–42
frictions 51, 56–7, 67, 72
Friedman, M. 90, 202, 204, 205, 208, 247, 250–51, 265, 314, 549, 567, 573
'fringe of unsatisfied borrowers' 135
full employment
definition 308
equilibrium 241, 334
fiscal policy
functional finance approach to 365–8
mainstream view 358–9
post-Keynesian view 364, 371
neoclassical theory 89–90, 311, 313, 325
potential negative effects for capitalists 296–7

prevailing, assumption 38, 85, 334
as threat to 'discipline in the factories' 609
traditional theory of comparative advantage 447
functional finance 365–8, 370, 595
functional income distribution 412, 546
fundamental economic table 93–7
fundamental prices 83, 85, 103
fundamental uncertainty 253, 389
funding of the deficit 369

G20 (Group of 20) 486
Galbraith, J.K. 296, 567, 572, 573, 575–6, 585
Garegnani, P.A. 77, 101, 104, 107
General Agreement on Tariffs and Trade (GATT) 454
general equilibrium 82–3, 87–9, 91, 106, 113, 274
general price level 161, 204, 333, 335, 336–7, 338–9
The General Theory of Employment, Interest and Money 28, 31, 101, 123, 145, 178, 193, 311, 328, 410, 423, 594
Geneva Conference (1947) 454
genuine progress indicator (GPI) 584
Germany 111, 145, 207, 367, 448–50, 453, 496–7, 501, 502–4, 508, 522, 588
Gertler, M. 184
Gini coefficient 415–16
Glass-Steagall Act (1933) 549
global demand 336, 337, 443
global financial crisis *see* financial crises
global power (dominant) 521
global supply 336
global trade imbalances 448–50
global value chains 462
The God Delusion 603
gold convertibility 497
gold standard (1879–1931) 497, 549, 559
Golden Age of capitalism 288
golden rule 344–6
goldsmith bankers 154
goods and services
ability to consume 377
aggregate demand and 307, 309–10, 311, 312, 313, 315–16
circular flow of 406–7
competitive prices of 476
consumer price index 339–40
consumption expenditures and 233
demand for 38, 45, 85, 208, 218, 234, 235, 270, 289, 334, 342, 407, 472, 479
exports and imports of 320, 390
factor substitutability and 271
foreign trade and 316
free flows of 454
GDP and 417, 582–3
government expenditures on 316, 319–20, 371
household spending on 322
income multiplier and 322
money and 337, 343, 345, 404, 406
non-price competitiveness and 470, 479–80
ongoing production of 419
output of 323–5, 582
in pessimistic times 33–4
price stability and 335
real analysis and 127
share of internationally traded 439
supply of 270, 307, 342, 407
unit labour costs and 501
world trade in 439–40
Goodwin, R.M. 297
Gossen, H.H. 86
government bonds 186, 189, 194, 195, 211, 369–70, 498–9, 500
government debt 179, 368–9, 370, 481
government deposits 167–8
government expenditure
automatic stabilizers and 363, 364
budget deficit and 355, 356, 362, 370, 546
categories 356
changes in balance sheets induced by 168
changes in policy mix 555–7
crowding in and crowding out 355, 361, 368
ECB and 498, 499
expansionary financial contraction and 29
fiscal policy and 235, 355, 356, 357
GDP and 35, 235
on goods and services 316, 319–20, 371
inflation and 334
Keynesian economics and 427–8, 481
in monetary circuit with state 142–3
MPC and 234, 243–4, 260
potential use to stabilize aggregate demand 194
quantitative easing 211
supermultiplier and 109–10
see also public expenditure
government policy 45, 412, 452, 456, 521, 555–6, 557
see also fiscal policy
government(s)
aggregate demand and 309–10
approaches to 32, 33, 34, 55

central bank independence and 221–2, 499–500
control over paper money 157–8
fiscal equalization mechanism 504
fiscal profligacy 367
functional finance and 366, 367–8, 595
interventions and market failures 45, 61
liquidity and 195
market stability 34, 145
neoliberalism and 599–600
as obstacle to market equilibrium 32
responses to economic crises 27–30, 43, 52–3
role in monetary circuit 142–3
wealth and 404
Graziani, A. 129, 130, 160, 172
Great Consensus 54, 58, 59
Great Crash (1929) 179
Great Depression 27–8, 30, 43, 110–11, 179, 308, 386, 454, 537, 545, 553
Great Moderation (1985–2007) 51, 53–4, 58–9, 61, 62
Great Recession 27–8, 54, 261, 364, 367, 386–7, 392, 529–30, 553, 615
see also financial crises
'Great Stagflation' era 598
The Great Transformation 43, 536
Greece 180, 323, 450–51, 495, 501, 503, 505
Greenwood, D.T. 567, 569, 572, 573–5, 581–2, 585
gross domestic product (GDP)
aggregate demand and 310, 312, 315
consumption and 234–5, 316
economic growth and 377, 378, 382
finding alternative measurement to 582–5
in international trade 320–21, 439, 448, 461
investment and 318
macroeconomic analysis 308–9
measuring foreign indebtedness 473–4
in neoclassical theory 313
and Keynesian theory 311
social product 92
United States 558, 574
growth *see* economic growth
growth of potential output 206, 383, 470, 472, 475–6, 482
growth rates
economies comparison 379–82, 471
of labour productivity 294
pre-Asian financial crisis 478
quantity theory of money 205–6
realized 112
secular stagnation and 29
United Kingdom 479
see also balance-of-payments constrained growth rate
Guide to Post-Keynesian Economics 227
Guttmann, R. 561

Haldane, A.G. 30, 66, 68–9, 72, 527, 531
Hamouda, O. 256
Hansen, A.H. 29, 30, 300
Harcourt, G.C. 103, 113, 294
Harmonized Index of Consumer Prices (HICP) 499
Harrod–Domar model 568
Harrod, H.R.F. 289, 290, 293, 303–4, 488, 568
Harrodian growth theory 387–9, 391, 396
Hartwick, J.M. 569, 570, 571
Hartwick–Solow rule 571
Harvard University 31, 55
Heckscher–Ohlin model 445, 463
hedge finance position 551
hedge funds 186–7, 195, 524
hedging 178
Hein, E. 521, 528, 530
helicopter money 169–70, 202, 216, 217
heterodox economics 3, 43–5, 172, 180, 284, 304, 490, 589, 601
absolute competitive advantage 448–50
aggregate demand 309
balance of payments constrained growth 470, 472–88
banks and modern payment system 161–9
central banking and monetary policy 213–24
consumption 253–60, 261
European Monetary Union 495, 502–9
financial system 185–90
financialization 518, 531
imbalances and crises 542–6, 547
international trade 447–51
investment 288–98
microfoundations
for heterodox macroeconomics 611–12
for macroeconomics critique 602–12
monetary economy of production 128–43
money and banking 156–61, 169–70
rejection of DSGE models 69
sustainability 569–70, 571, 572–3, 576, 579, 581–2, 585
true economic development and 575–9
unemployment 447, 450–51

wealth distribution 419–27
see also post-Keynesian economics
heterogeneous firms 446
hierarchical reductionism 602, 604
The High Price of Bullion, a Proof of the Depreciation of Bank Notes 118
Hilferding, R. 190, 191–2, 194, 519, 520
historical growth record 378–9, 381–2, 392, 396
history 30, 43, 71, 82, 396, 543, 597
of economics 84–92
further reading 114–15
of political economy 92–113
hoarded savings 122, 138, 141, 343
hoarding 86, 96, 141, 337, 339
Holt, R.P.F. 227, 567, 569–78, 581–2, 585
'*Homo oeconomicus*' hypothesis 238, 250, 253, 254
household savings
appearance in mainstream theory 182
banking and 155–6, 158–61
behaviour 186, 388–9
disposable income and 318, 360
equations for 387–9
firms and 138–9, 141, 193
interest rates and 274
loanable funds and 274–5
outlets for 369–70
see also financial savings; hoarded savings
household(s)
as borrowers and lenders 183
consumption and 233, 243, 249–50, 256–9, 261, 280, 283, 300, 309, 313, 316–17
crowding out 361
financialization and 527–8
inflation and 339, 598
marginal propensity to consume 322
in money triangle 215, 216
purchasing power 338
reflux principle and 137–8
utility maximizing 82
wages and 544, 546
wealth distribution 405–6
How Much is Enough? The Love of Money and the Case for the Good Life 615
human capital 269, 531, 573–4, 581, 590
Human Development Index (HDI) 584–5, 589, 590
human nature 42, 85, 240, 243
Hume, D. 48, 204
hypotheses 34–5, 36, 310–11
hysteresis 71
ideology 31–4, 64, 100, 526, 537, 550, 575
imbalanced trade 447, 450
imbalances
cyclical growth dynamics 541, 547–53, 563
in EU 502–3
external adjustment 540–41, 553–60
growth and distribution 546–7
heterodox perspective 542–6, 547
mainstream view of self-adjusting economy 541–2
nature of 541
in private sector, correction of 326, 327
social 585
see also financial crises; trade imbalances
immutable laws 35
imperfect competition 76, 92, 466
import demand function 470, 472, 476, 482
import income elasticity of demand 470, 475–88
import price elasticity of demand 475–6
imports
aggregate demand and 309–10, 320–21
and balance-of-payments constrained growth rate 470, 472–88
consumer price index 339
currency depreciation 326
European rebalancing mechanism 508
income multiplier and 322–3
in Kaldorian growth theory 390–91
in Kalecki's profit equation 292
policy mix changes 555–7
supermultiplier and 110
see also international trade
income
circular flow of 390, 406, 610
consumption and 234–5, 236, 238, 240–52, 255–61
divergence 392, 395
full employment 358–9, 367
growth in Latin American and Asian economies 378–81
investment and savings linked with 358–9
in Keynesian growth theory 392–3
money and 343–6
see also disposable income; national income; net income; per capita income
income channel 218–19
income distribution
aggregate demand 530
classical 420–21
conflict over 41
consumption and 258, 317

definition 403
economic growth and 41, 377, 392–3, 398, 466, 546–7, 560, 596
financial investment and 278
financialization and 532–3
formation of prices in relation to 107–8
functional 412, 546
inflation and 335, 346
interest rates and 412–13
Keynesian monetarist model 412–13
Keynes's theory 413, 426
marginal productivity and 407–11
microeconomics approach 407–11, 419–20
monetary policy and 218–21, 225
output and employment 111
states of 413–19
theoretical approaches to 82–3, 99, 118, 412–13, 419–20, 422, 423, 426, 430, 466
variables affecting determination of 424
wealth and 403–4
workers and firms 335
income elasticity of demand for income and exports 470, 475–88
income flows 154, 155–6, 159–60, 197, 218
income inequality 43, 376, 392–3, 534
income multiplier 308, 321–3, 325, 326, 328, 596
elaborated 322–3
simple 321–2, 323
income redistribution 244, 393, 403, 404–5
Income, Saving and the Theory of Consumer Behavior 256, 263
increasing risk, principle of 293, 373
indebtedness 192, 197, 273, 274, 293–4, 346, 428, 473–4, 551
Index of Human Development (IHD) 418
Index of Sustainable Economic Welfare (ISEW) 584
India 180, 416, 445, 448–9, 452, 521
indirect labour 81, 97, 98, 99, 107–8, 422
indirect taxation 338
individual behaviour 27, 45, 213, 597, 603–4, 608–9
Indonesia 448–9, 478
industrial capitalism 189, 520
Industrial Crises in Contemporary England 563
inequality 108, 413, 414–18, 439, 446, 456, 532, 574, 582
see also income inequality
infant industry protection 438, 452–3, 460
inflation
agents' behaviour and 339
causes 334–5
central bank independence and 221–2
cost of living and 339–40
definitions 332, 333, 335–8
Great Moderation era 54
measurement 333–4
as 'monetary phenomenon' 205
monetary policy
ECB 499, 500–501
heterodox perspective 213, 214, 220–24
mainstream perspective 202, 203–9
money issuance examples 344–6
neoclassical view 90
price level and 338–9
remaining steps leading to full explanation of 350–51
stagflation period (1970s) 90, 207, 386, 399, 481, 547, 598
towards monetary macroeconomic analysis of 343–50
traditional analysis 333–40
criticisms 335–8
Inflation, chômage et malformations du capital 352
inflation rates 53, 206–9, 220, 498, 552, 598
inflation targeting 205, 208, 220–22, 225
inflationary expectations 598
initial finance 159, 160
An Inquiry into the Nature and Causes of the Wealth of Nations see Wealth of Nations
insolvency 185
instability
banking 183
capitalism's long record of 541
Harrod on 303–4
investment 279, 289–90, 423–4
market 32–3
see also financial instability
institutional change 35–6, 545, 575
institutionalists 575–6
institutionalized scarcity 579
institutions 27, 32, 37, 46, 52, 69, 92
banking 153, 155, 157, 169
constituting financial markets 160
financial innovation and 524
global 486
see also financial institutions
insurance companies 180, 186–7, 194–5
interbank market(s) 164, 185, 211, 212, 215–17
interest, definition 177

interest payments 357, 362, 369, 371, 372
interest rate channels 208, 210, 212
interest rate rule 208, 220, 223
interest rates
central banks and 206, 208–11, 212, 217–18, 225, 360, 398
consumption and 213, 236, 238, 241, 249
ECB monetary policy 499, 500–501, 508
GDP growth rate and 370–71, 372
indebtedness and 293
investment and 214, 271, 273–4, 275–6, 300, 358–9, 426, 428
Keynesian policies contributing to high 59
Keynes's use of monetary policy to drive down 194
low
failing to stimulate economic activity 211
as generating asset bubbles 219
government subsidy of industries 452
as policy to deal with financial crises 52, 180, 195, 550
natural 223–4, 225, 358–9
nominal
long-term 498
zero lower bound for 136
overnight 168
policies affecting income and wealth distribution 220, 223, 224
post-Keynesian perspective 214–15
raising in balance-of-payments crisis 474
textbook versus real world approaches 187–8
unemployment and 360
intermediaries/intermediation 67, 151, 154–6, 159, 160, 172, 179, 185–8, 344, 346, 549
international banking system 183
International Clearing Union 464
international crisis dynamics 553–4
'international euro' 508–9
international monetary architecture 506
International Monetary Fund (IMF) 63, 71, 146, 443, 449, 478, 479, 486, 497, 508, 512
international monetary system 459–60, 462, 467, 512, 559–60
international payments 352, 504, 512
international settlement institutions 506, 507, 508
international trade
aggregate demand 320–21
heterodox perspective 447–51
importance 438–9
Kaldorian growth theory 390–91
liberalization, trade agreements and trade wars 453–60
long-run development and infant-industry protection 451–3
mainstream perspective 440–47
manufactured exports, fallacy of composition and global value chains 460–62
non-price competitiveness 487–8
recommendations for 462–3
winners and losers 445–7, 462
see also exports; imports
intertemporal choice 60, 234, 235, 237–8, 241, 247, 250, 252
investment
capital 522, 569, 571, 573, 575, 590
consumption and 308, 309–10, 315, 373, 405
decisions 42, 140, 186, 272, 276–7, 288, 294–6, 298, 360, 530, 550
determination of long-period or trend output and employment 109–10
different types of 269
empirical testing of functions 297–300
employment and 190–94, 278, 297, 318, 424–6, 533
expenditures 101, 154, 355, 362, 486
financialization and 299–301, 530–31
heterodox perspective 288–98
importance 269–71
instability 279, 289–90, 423–4
intentions 357, 360, 366, 367–8, 369
Keynes's approach to 275–88, 423–6, 428, 598
mainstream view 141, 271–5
planning of 139–41
profit and 276, 291, 293, 294–8, 318–19, 347–51, 352, 360, 388–9, 547–8
requirements and preferences 188
savings and 189, 275, 290, 291–3, 313–14, 315, 318, 327, 357–60, 366, 367–8, 412–13, 519, 556, 608, 610
spending 274, 280, 281, 287, 289, 292, 301, 386, 389–90, 393, 548
for sustainable development 579–82
trade and 456–7, 462
see also mal-investment; overinvestment; private investment; public investment; underinvestment
investment-driven growth 387–9
investment multiplier 81, 244, 281–3, 304
investment rate 193, 292, 304, 471, 472

invisible hand 48, 541
involuntary unemployment 81, 83, 90–91, 96, 101, 108–11, 275, 313, 332, 341–2, 343, 600
Ireland 449, 450, 455, 502–3
Italy 129, 367, 449, 450, 478–9, 503

Jackson Hole conferences 61, 62
Japan 111, 367, 378, 439, 448–9, 453, 455, 457, 471, 481, 522, 595
Jespersen, J. 316
Jevons, S. 87, 125, 239, 407, 541, 596
Jorgenson, D.W. 271–2
Journal of Post Keynesian Economics 227

Kahn, R. 76, 307, 321, 327, 328
Kaldor, N. 71, 270, 398–400, 488, 490, 579
Kaldor–Verdoorn Law 270, 394
Kaldorian growth theory 390–91, 394–5, 396, 398–9
Kalecki, M. 191–2, 194, 293, 296, 317, 326, 365, 370, 373, 530, 546–8, 560, 608, 609
Kaleckian growth theory 389–90, 391, 393, 396
Kaleckian paradox of costs 607
Kalecki's profit equation 291, 292, 607
Keynes, J.M.
  absolute income hypothesis 242, 245, 257–8
  analysis of finance 193–4
  biography *see The Return of the Master*
  borrowers' and lenders' risk 293
  business cycles 548, 551
  on capitalism's propensity for crisis 544–5
  versus 'Classics' 189
  consumption approach 238–44, 246, 247, 250, 256
  criticism of mainstream theory 238–40
  economic growth and well-being 567
  efficiency wage 547
  on the future 222, 577
  on hedging assets 178
  insight into imbalances and crises 96, 544–5
  International Clearing Union 464
  on investment 110
  investment approach 275–88, 423–6, 428, 598
  investment multiplier 81
  model 419
  multipliers 104, 283, 322
  paradox of thrift 84
  plan 506–7, 508, 512
  on policy rate of interest 210–11
  on population growth and natural resource capacities 577
  portrait of 145–6
  principles of radical uncertainty 577
  quotes from 25–6, 28, 31, 37, 39, 40, 42, 66, 82, 115, 123, 128, 129–30, 135, 139, 143, 201, 215, 279, 303–4, 319, 506–7, 594, 609
  on relative real wage 255, 256, 263
  reliance on principles 107
  significance of consumption 240–41
  social nature of decisions 263
  uncertainty approach 253
  voluntary and involuntary unemployment 340–42
  see also A Treatise on Money; animal spirits; bancor; effective demand; liquidity preference (LP); marginal efficiency of capital (MEC); marginal propensity to consume (MPC); *The General Theory of Employment, Interest and Money*; underconsumption
Keynesian consumption function 234, 241–4, 596
Keynesian economics
  aggregate demand 311
    economic growth and 376, 384–6, 387
    macroeconomics approach 387, 411–13, 426
    and output 234, 311, 315–21
    policy guidance 427–8
    and supply 423–4, 426
  balance-of-payments 474–5, 481
  case for government intervention 253
  consumption equation 555–6
  demand-led growth 376, 384–6
  fallacy of composition examples 607–8
  fiscal stimulus 28, 58
  hedging assets 178
  as increasingly discredited 498
  monetarism as direct challenge to 205
  neoclassical synthesis 89–90
  rejection of Keynesian-inspired fiscal policies 58–9, 235
  theories on determinants of investment 283–4
  theory of output and employment 109–11, 114, 234
    'as a whole' 594, 597
  uncertainty 239–40
  see also classical–Keynesian political economy; new Keynesian economics
Keynesian growth theory 383–4, 387–91
  properties of 392–6
Keynesian models 68, 103, 244, 403, 411, 498
Keynesian Revolution 101, 111, 303, 545–6
King, J.E. 56, 63, 67, 254–5

Krippner, G. 519
Krugman, P.R. 53–5, 62–3, 65–6, 68, 70, 214
Kuznets, S. 245–7, 255, 258, 352, 389

labour
  aggregate demand and 309, 310, 311–15
  agricultural 94–7, 99
  commodification of 536–7
  comparative advantage and 440–42
  demand-led growth 385
  direct and indirect 81, 97, 98, 99, 107–8, 422
  in economic growth model 569
  effective demand and 323–5
  financialization and 527, 533
  in formation of prices 107–8
  hiring decisions 132
  income distribution and
    heterodox perspectives 419, 420–22
    macroeconomics approach 411, 412–13
    microeconomics approach 407–9
    new macroeconomics approach 423–4, 427
  international trade and 445–7, 448–9, 451, 454, 458–9, 460, 462
  investment and 271, 284, 298
  long-run growth rate and 472
  luxury consumption and 191
  marginal product of 104
  as natural factor of production 104
  supply-led growth and 382–4
  values 83, 97–8, 100, 106–7
labour costs 98, 108, 276, 423, 440–41, 448, 501, 545, 547
labour markets
  aggregate demand 324–5, 411, 412–13
  changes in rate of interest 218–19
  economic theories 87–92
  full employment 326, 609
  inflation and 214
  monetary policy and income distribution 225
  'monomani' in 427
  neoclassical 312–15
  'reserve army' of unemployed 44
  unemployment as consequence of rigidities in 342–3
  wage determination 544
labour theory of value 97, 407, 563
  classical 403, 422
labour value principle 81, 98, 103, 105, 107–8, 118
Lafleur, L.-R. 135
laissez-faire approach 31, 537
Largentaye, J. de 172
Laspeyres price index 333–4, 338, 339
Lavoie, M. 54, 70, 130, 136, 139, 220, 227, 254–5, 295–6, 304, 578
law of diminishing marginal utility 86
law of equalization of the marginal utility–price ratios 86
'law of one price' 477
laws of behaviour 60
'The laws of return under competitive conditions' 76
'LDC debt' crisis 553–4
learning by doing 394, 471
Lehman Brothers 62, 179, 180, 201, 212, 503, 552
Lerner, A. 365–6, 476, 561
life-cycle hypothesis of consumption function 234, 247–50, 260–61
life-cycle theory 250
Lima, G.T. 480
Lionel Robbins lectures 63
liquid savings 138–9, 141–2, 154, 160
liquidity 183, 187–8, 189, 195, 212, 279, 424, 427, 529, 607, 608
liquidity constraints 249–50, 252, 326
liquidity preference (LP) 212, 275–6, 360, 403, 410–11, 426
loanable funds 160, 189, 274–5, 549
loanable funds approach 358, 360
loanable funds market 155–6, 274–6
long-period Kondratiev cycles 111
long-period theory of output and employment 109–10, 114
long-period trend unemployment 112
long-run aggregate supply (LRAS) 314–15
long-run consumer behaviour 245–6
long-run consumption function 246, 252, 257–9
long-run development (international trade) 451–3, 462
long-run economic growth 378–9
  model 598
  rate of 71, 205, 313, 469, 472, 488
long-run equilibrium condition 290
long-run sources of aggregate demand 386–7
long-term finance
  easier access to and provision of 189–90, 195
  economic consequences of 190–93
Lorenz Curve 414–18
Lucas, R.E. 53–4, 55, 58, 59–60, 63, 91, 567, 594, 598–9
luxury consumption 191

M0-M3 money supply measures 203, 204
Maastricht criteria 495, 498–9
The Macrodynamics of Advanced Market Economies 227
macroeconomic equilibrium 80, 81, 104, 110, 387–90
macroeconomic models 244, 261, 271, 316, 321, 529, 530
macroeconomic policy 59, 595, 598, 599
macroeconomic stabilization 224, 490
macroeconomic theories
  Keynesian 311, 387, 413
  neoclassical 311, 313
  of production, employment and aggregate demand 311, 313
macroeconomics 51–72
  approach to aggregate demand 411–13
  Caballero on 114
  crises facing modern 51–3
  disappearance of 594
  dissatisfaction with current state of 63–4
  heterodox 611–12
  mainstream perspectives 45, 58–62, 70, 595–602
  metaphors 610–11
  microeconomics and 45–6, 594, 598–9, 601, 610–13
  microfoundations and 51, 55–6, 58, 60–63, 67–70, 255, 611–12
  modern versus Keynesian 595–6
  new approach to stock of wealth and social well-being 423–7
  post-Keynesian perspective 69–72
  pre-crises era 53–8
  rate of interest as key variable in 187
  reconstructing 66–72
  rethinking models 62–6
  see also DSGE (dynamic stochastic general equilibrium) models; New Consensus Macroeconomics
macrofoundations 608–9
mainstream economics
  balance-of-payments constrained growth rate 470–71
  banks/banking 151, 152–6
  central banking and monetary policy 202–13, 224–5, 335
  civil war 51–2, 53, 55
  consumption function 235–8
  consumption theories 247–53, 259–60, 261
  as in disarray 52–3
  dissatisfaction with current state of 54
  economic growth (supply-led) 382–4, 470–71
  European Monetary Union 495–502
  financial system 181–5, 189
  fiscal policy 358–9, 361–2, 371
  inability to predict financial crisis 38–9, 52, 518
  international trade 440–47
  investment 141, 271–5
  Keynes's criticism of 238–40
  macroeconomics 45, 58–62, 70, 595–602
  microfoundations dogma 595–602
  monetary economy of production 123–8
  money and banking 152–6, 161–2, 169–70, 172
  non-acceptance of scientific approach 37
  self-adjusting economy 541–2
  sustainability 567, 569–71, 572–3, 583
  wealth distribution 405–13
  see also neoclassical economics
mal-investment 550
managerial capitalism 296, 298
Mankiw, N.G. 31, 55, 599
manufactured exports 460–62
marble and bowl analogy 32, 33
marginal analysis 86–7
marginal costs 76, 81, 86–9, 101
marginal efficiency of capital (MEC) 276, 403, 410, 425, 426
marginal principle 81, 98, 100
marginal productivity 81, 86–7, 104, 276, 407–9, 412, 420–21, 425, 569
marginal propensity to consume (MPC) 242–6, 249, 259–60, 321–3, 403, 410–11, 426–8
marginal utility 81, 86, 87–8, 408
Marginalist Revolution (1870s) 85, 87, 100, 125, 541, 560, 596
market equilibrium 83, 90, 290, 314, 544
market mechanism 213, 215, 422, 541–2, 550
Marshall, A. 25, 76, 82, 83, 87–9, 100–101, 118, 541–2, 544
Marshall–Lerner condition 476, 561
Marshallian cross 87–8
Marx, K.H. 40, 48, 77, 96, 102, 118, 192, 256, 342, 546, 548
  on declining profitability 543–4
  Marxian perspectives 108, 297, 422, 520, 563
  portrait of 430–31
  reactions against 99–100
master economists 25–6
Mathematical Investigations in the Theory of Value and Prices 241

mathematics
parallel lines 66
risk and 188
use of 37, 39–40, 254, 542
maturity transformation 187–8
McCombie, J.S.L. 472, 479, 487
means of payment 152, 156–7, 170, 504, 507, 512
The Megacorp and Oligopoly 227
megacorps 296
Menger, C. 87, 407, 541, 596
Mercosur 454
Merkel, A. 595
'mesoeconomics' 611
metaphors 20, 283, 593, 594, 601, 603, 609–11, 613
methodological individualism 596–7
methodology 594, 599, 608, 613
Mexico 183, 377, 381, 446, 448–9, 454–5, 458–9, 461–2
micro-reduction 594, 597, 601, 603, 604, 613
microeconomics
approach to income distribution 407–11, 419–20
macroeconomics and 45–6, 594, 598–9, 601, 610–13
macrofoundations for 608–9
metaphors 611
rational consumer and 239, 247
microfoundations
acceptance of
disappearance of macroeconomics 594
implications for economic policy 595
arguments in favour of 600–602
definition 51, 594
DSGE models and 55–6
economists' debate on 55–6, 58, 62–3, 67–8, 69–70, 261
heterodox critique 602–12
importance of concept 612–13
macroeconomics and 58, 60–63, 67–70, 255, 594–5, 611–12
mainstream perspective 595–602
metaphor 609–10
origins of dogma 596–600
RARE 600, 609
transforming economics 61
microfoundations delusion 62–3, 600, 613
Miller, M. 182
Minsky, H.P. 71, 293–4, 518, 553
financial instability hypothesis 70, 197, 551–2
portrait of 197
'Minsky moment' 552–3
models/modelling 35, 36, 52–3, 70–72, 98, 106, 598, 601, 604
'correct' 221
criticisms 62–6, 68–9
mainstream perspective 58, 60–62
and policy 57
post-Keynesian perspective 69
rejecting 66–9
risk 188
scientific realism and 610
use of 37–40
see also DSGE (dynamic stochastic general equilibrium) models
modern capitalism 36, 189–90, 191, 438
modern economies 53–4, 133, 234, 448
modern finance 188–9, 194–5
modern macroeconomics 595–6
Modern Monetary Theory (MMT) 172
modern payments system 163–6
Modern Socialism in Its Historic Development 563
modern welfare states 308, 322–3
Modigliani, F. 182, 247, 264–5
Modigliani–Miller theorem 182
monetarism 205, 241, 366, 411
as abandoned by central banks 206
collapse of 207
core idea of 208
and new classical economics 90–91
monetarist economics
aggregate demand 411–13, 426
mantra of 205
monetary union 496–7
monetary base 202–3, 204, 205–6, 369
monetary circuit
banks and 133, 134, 143, 152, 159, 172
definition 152
role of state in 141–3
theory of 129–30, 143
core assumptions 130–31
implications 141
post-Keynesian 123, 352
stages of 131–41
Monetary Economies of Production: Banking and Financial Circuits and the Role of the State 172
monetary economy 128, 129–30, 201, 427
monetary economy of production
based on capital accumulation 341
definition 122
heterodox perspective 128–43

modern economies best described as 448
neoclassical/mainstream perspective 123–8
Monetary History of the United States 1867–1960 205
monetary policy 200–25
central banks 274, 319, 335, 557
credibility 221–2
demand management policies 307–8
ECB 499–500, 502
heterodox perspective 213–24
instruments 326
mainstream perspective 202–13, 224–5, 335
from post-Keynesian perspective 214–24, 225
cyclical 550, 552
income distributive nature of 218–21, 225
Keynes and 194, 201, 210–11, 426
turn towards 60
'monetary policy dominance' 59
monetary production economy 80, 92, 103, 107–8, 113, 114, 201
monetary targeting 201, 202–3, 206–8, 220
monetary union
approaches to 496
definition 495
implementation factors 498
viable 503
see also European Monetary Union (EMU)
money
and banking
balance sheet perspective 161–9
heterodox perspective 156–70
mainstream view 152–6, 161–2, 169–70, 172
crucial importance of 83–4
destruction of 137–9
income and 343–6
as means of payment 151, 156–7, 170
nature of 215–18
place in textbooks 124, 152–3, 161, 169–70
purchasing power of 332, 335–6, 342, 344, 346, 350, 351, 506
relationship with barter 123–4, 126, 128, 153, 157, 170
as store of value 103, 204
theory of 127, 128, 129, 134, 156 *see also* barter theory of money; quantity theory of money
'velocity of circulation' of 204–5, 206
Money and the Mechanism of Exchange 125
money creation
bank credit and 132–4, 142, 162, 170, 549
ex nihilo 158–9, 162, 169, 172
mainstream story of 161–2
money multiplier 155, 161, 201, 202–3, 217
money supply
central bank control over 202, 203, 209, 399, 549
as credit-driven and demand-determined 200
inflation and 203–6, 208, 223–4, 334, 335, 337, 399
measurements of 204
monetarism and 205
money multiplier and 201, 203
multiple expansion of 155
total 154–5
money zero maturity (MZM) 204
Monnaie et Macroéconomie: Théorie de la Monnaie en Déséquilibre 172
mortgage-backed securities (MBSs) 525
MPC *see* marginal propensity to consume (MPC)
multinational corporations (MNCs) 180, 452–3, 455–6, 461
multiplier principle 119, 268, 281–2, 285–6, 322
multipliers
fiscal 235, 244, 260–61
importance of MPC for 244
tax 234, 243–4
see also income multiplier; investment multiplier; money multiplier; private investment multiplier; supermultiplier

national accounting 310, 419
National Accounting System (NAS) 308–9, 310–11, 318–19
national central banks (NCBs) 497, 504, 506, 507–8, 512
national currency 335–6, 339, 497, 506
national income 81, 96, 177, 244, 264, 286, 292, 300, 357, 414, 442, 533, 555–6
nationalism *see* new economic nationalism
natural capital 570–73, 574, 576–7, 579–81, 582, 590
natural employment 312–13
natural factors of production 104
natural prices 85
natural rate of growth 110, 205, 290, 470, 490
natural rate of interest 223–4, 225, 358–9
natural rate of unemployment 314, 490
natural resources 438, 445, 460, 567, 569–73, 576–7, 582
'natural' unemployment 308, 314, 490
natural wage rate 99
necessary consumption goods 93–5, 99, 108

negative savings 187, 236, 248–9
neoclassical economics
aggregate demand 307, 311, 312–14
aggregate supply model 314–15
belief that markets are self-stabilizing 145
capital 275
causes of inflation 334–5
classical value theory 100
economic growth
balance of payments constraint 477
and development model 567, 568–9, 589, 598
equilibrium rate of growth 398
models 270
as supply-led 377, 382–3
trajectories 385
equilibrium models 105–6
exogenous shocks 542
factors of production 104, 297–8, 311
full employment 311, 313, 325, 342
Heckscher–Ohlin model of trade 445
income distribution 405–6, 408–9, 412
internal critique of 114
investment theories in 271, 272–3, 275, 276
market failures 45
natural sciences flavour of 111
origins of money as linked to barter 123–8, 143
pre-World War II dominance 422
price theory of supply and demand 227
pure theory of price 118
putting individual at heart of analysis 130
Say's Law 132, 133
supporting 'putty' models 284
sustainable development 570–79, 585
theory of value 403, 407
voluntary unemployment 308
see also mainstream economics
neoclassical–Marshallian theory 88–9
neoclassical model of economic growth and development 567, 568–9, 589, 598
neoclassical school 87–9
counter-revolution against 100–102
neoclassical synthesis 89–90
Keynesian 55, 411
new 70
neoclassical theory of value 403, 407
neoclassical–Walrasian economics 83, 87–8, 102, 103, 113–14
neoliberalism 71, 459, 518, 521, 537, 599–601
net debt 177
net income 526
net stock of foreign liabilities (NFL) 474
net worth 184, 249–50
Netherlands 320, 323, 444, 448–9, 522
network theory 185
new classical economics 91, 599
New Classical theory 182–3, 599
New Consensus Macroeconomics 208–9, 223, 499
new economic nationalism 456–60
A New Guide to Post Keynesian Economics 227
new Keynesian economics 91–2
new Keynesians 184, 274
new macroeconomics approach 423–7
new neoclassical synthesis 70
new technologies 270, 300, 318, 323, 524, 571
nominal exchange rate 473, 476–7
nominal interest rates 136, 359
nominal wages 209, 342–3, 449–50, 476, 547
non-convex adjustment costs 272
non-financial corporations
banks lending to 211
financialization and 522, 523–4, 526–7, 531
non-market values 583
non-performing loans 169, 212
non-price competitiveness 470, 479–80, 482–3, 485, 487–8
North American Free Trade Agreement (NAFTA) 454–8, 461
Norway 416, 449
Not in Our Genes 604

offshoring 179, 446
opportunity costs 273, 274, 438, 440–41, 447
orthodox economics *see* mainstream economics; neoclassical economics
oscillator model 268, 285–8
output
aggregate demand
Keynesian theory of 234, 315–21
neoclassical theory of 312–15
in agricultural sector 36, 85, 90–91, 95–6, 99
balance-of-payments constrained growth rate 470–71, 472, 474, 481–2, 484–6, 488
in comparative advantage theory 440–41
consumer price index and 339–40
employment and 90–91, 96, 101–2, 105, 107, 109–13, 114, 190–93, 221–4, 315, 594, 597
exchange rate pass-through and 210
gaps 201, 223

of goods and services 323–5, 582
in income distribution theory 405–6, 409–10, 412, 423
income multiplier and 322
inflation and 334–5, 336
investment and 360, 388–91
in Keynesian growth theory 393–5
marginal productivity and 81, 86–7
money and 128, 129, 130, 134, 137, 205–6, 336, 337–8, 344–6
supply factors and 88–9
sustainable development and 570, 574, 575, 582
trend 109–13
volatility 54, 59
see also potential output
over-the-counter (OTC) derivatives 528
overinvestment 296, 527
overnight market(s) 164–5
see also interbank market(s)
overproduction 81, 86, 101, 275, 540, 543–4, 546, 547, 552, 560
Oxford Review of Economic Policy 55, 66–7

paradox of costs 607
paradox of liquidity 607
paradox of thrift 46, 84, 138, 293, 318, 392, 396, 607, 608
paradox of tranquillity 294
'Pareto optimum' 82–3
Parguez, A. 129, 137, 157, 172
partial equilibrium 76, 87–9
path dependence 395–6, 579
'pattern of accumulation' 519
payments
debt service 197
euro area 505–7
factor 409, 420–22
interest 357, 362, 369, 371, 372
international 352, 504, 512
payments and settlements systems
banks/banking 163–5
definition 152
modern 161–9
per capita income 377–82, 393, 395, 444, 567, 583–4
perfect competition 56, 76, 88–9, 312, 314, 325, 612
permanent income hypothesis 234, 247, 250–52
Phillips curve 209, 213, 214, 225, 598
PIIGS countries 503
Piketty, T. 574, 605, 616
pluralism 30–31, 579
Polanyi, K. 43, 536–7
policy
changes in mix 553–60
imbalances 555–7
and models 57
policy-constrained countries 481
political economy
approach 83–4
history of 92–113
see also classical–Keynesian political economy
political science 26, 42
Ponzi finance 197, 551–2
Popper, K. 597
portfolio theory 181
Portugal 450, 503
post-Keynesian economics
accounting decomposition of profit rate 299
aggregate demand 133
demand-led growth 598
financial crises 551, 552
fiscal policy 360–62, 365, 371
Harrod–Domar model 568
investment and 112, 280, 298–9
monetary policy 214–24, 225
principle of hysteresis 71
radical uncertainty 246–7, 577–8
on state of macroeconomics 69
sustainable development 572, 577–9
see also heterodox economics
post-Keynesian sustainability model 577
post-Keynesian theory
of consumer choice 578
of economic growth 398
of the firm 294–6, 298
of the monetary circuit 123, 352
of the profit rate 398
potential output 191, 206, 222–3, 312, 325, 383–6, 394–5, 470, 472, 482
Poverty and Famines: An Essay on Entitlement and Deprivation 589
pretense-of-knowledge syndrome 114, 222
price competitiveness 469, 475–7, 479, 488
price level
falling 595, 612
heterodox perspective 161–2
inflation and 333, 334–5, 336–7, 338–9
mainstream perspective 202, 204–5, 208–10

monetarist perspective 90
monetary targeting 208
in neoclassical aggregate supply 314–15
post-Keynesian perspective 214, 217–18
in quantity theory of money 204
price setting 545
prices
absolute 83
asset 172, 185, 200, 219
bond 184, 193
business cycles and 105, 548
capital 271
commodity 437, 443–4, 463
consumer 214–15, 499
domestic 476–7
equilibrium 87
factor 113, 284, 297–8
falling 545
of final goods 84
flexible 90
formation of 107–8
of imported goods 339–40
income distribution and 408–9, 412, 422, 426
inflation and 335–8
market 85, 101, 310, 334, 573
of production 83, 93, 97, 100–101, 103–4, 106–7
share 178, 527
sticky 490
stock 181, 184, 279, 524, 526
stock market 186, 278
see also relative prices
principle of effective demand 81, 84, 90, 101–5, 109, 373, 398, 596, 598
principle of propriety 84–5
principle of separability of needs 255
principle of supply and demand 89, 104, 118
principles
applied theory and 105, 106–7
pure theory and 105–7, 118
Principles of Economics (Marshall) 82, 89, 100, 118
On the Principles of Political Economy and Taxation (Ricardo) 77, 82, 98, 100, 118
Principles of Political Economy (Tugan-Baranovsky) 563
private consumption 308, 309–10, 316–17, 318, 320, 321–2, 326
private investment 59, 279, 288, 308, 315, 317–19, 326, 327, 357, 370, 474
private investment multiplier 234, 244
private sector
aggregate demand and 309, 315
bank transactions 162–6
debt 133, 179
firms/companies 138
governments and 34, 52, 142, 326–7, 357, 368, 369, 557
inflation and 334
investment 181, 189, 360
money multiplier and 203
relationship with public sector 357
savings intentions 369
spending 29
production
aggregate demand and 309, 310, 311, 313, 315, 316, 320–21, 323, 326
banks and 152, 158, 160, 170
in financial system 182, 186, 188–9, 191–3
financialization and 521, 527
fixed capital formation and 347–50
heterodox perspective 44–5
in history of economic theories 92–108
imbalances and crises 543–5, 546, 548, 553, 556
income distribution and 405–6, 408–9, 411, 419–25, 428–9
international trade and 320–21, 439, 445, 447, 450, 451, 461–2, 463
investment and 270, 271, 275, 282–7, 297–8
macroeconomic theories 311, 313
marginal costs of 88–9
money and 83–4, 342–6, 351
as motivated by profit 609
and natural capital 580–81
planning of 131–2
prices of 83, 93, 97, 100–101, 103–4, 106–7
purchasing power and 335–6, 352
putty and clay models 284
social and circular process of 83, 93–4
social nature of 92, 96
see also factors of production; monetary economy of production; overproduction
Production of Commodities by Means of Commodities (*PCMC*) 77, 100
production possibility frontier (PPF) 383–5
productivity
absolute advantage 441
of factors of production 377, 382–5, 394
income distribution and 425–6
international trade and 448–50, 460, 463

in Keynesian growth theory 392–5
labour 108, 110, 221, 325, 394, 460
long-run growth rate of output and 313
technological capability and 311, 341
see also marginal productivity
productivity growth
determining growth of standard of living 470
positive impact on export demand 399
rate 394, 471
Verdoorn Law 270, 394, 472
wages rates and 547
profit equation (Kalecki) 292, 607
profit maximization
of banks 154–6
behaviour 312
firms/companies 82, 271, 325
Walrus's basic model 87, 88
profit rates
accounting decomposition of 299
falling 552
in history of economic theories 100–101, 104, 107–8, 112–13
income distribution and 431
Marx on 543
natural 85
and quantity of capital 466
socially appropriate 83
theory of 398
profit-seeking behaviour 153–4, 157
profit(s)
financial
origin of 532
in US and UK 522
in financial system 189–93, 194
income distribution and 405, 419–21, 423–4
investment and 276, 291, 293, 294–8, 318–19, 347–51, 352, 360, 388–9, 547–8
production motivated by 609
public deficit and 292
Ricardo's theory of *see* corn model (Ricardo)
shares 546–7
short-termism and 527
wages and 308–10, 335, 392–3, 608
progressive tax system 363, 365
protectionism 453–4, 537
psychology 42–3, 605–6
public consumption 309
public debt 52–3, 221, 357, 362, 367, 372, 495, 498
public deficits 28–9, 32, 288, 292, 344, 359, 366–7, 369–71, 372, 498
public expenditure
austerity/fiscal consolidation 51, 59, 387, 495
crowding out effect 315, 361–2
expansionary fiscal contraction and 29
fiscal policy and 326
in Kalecki's profit equation 292
in monetary circuit with state 142
responses to financial crises 28, 34
sustainability and 573
see also government expenditure
public investment 315, 327, 356–7, 362, 457, 575
public sector
bank transactions 166–9
budget balance 221, 556–7
crisis interventions 212, 558
deficits 369, 387, 500, 546, 557
employment 315, 368
inflation and 334
intervention in market mechanism 213
lower rates of interest and 326
in monetary economy of production 142
relationship with private sector 357
sovereign debt crisis 501
public spending *see* public expenditure
purchasing power
of commodity exporters 432–3
consumer price index and 339–40
impact on exports 320
of income 431, 584
of money 332, 335–6, 338, 342, 344, 346, 350, 351, 506
need for income policies to enhance 107
scenario of declining 112
pure theory
of business cycle 111–12
of output 109
principles and 105–7, 118
putty and clay models 284

quantitative easing (QE) 172, 181, 195, 201, 211–12, 326
quantity theory of money 203, 204, 205, 208, 241, 335, 337, 399
Quesnay, F. 48, 93–7

radical authors 298, 299
radical theorists 297

radical uncertainty 270, 276–7, 577–8
Rajan, Raghuram 61–2
random walk consumption function 238
random walk theory of consumption 252–3
rational behaviour 60, 69, 80, 82–3, 87–8, 289
rational expectations 67, 221, 252–3, 255, 261, 599, 600, 601
rationality
  bounded 253, 254
  economic 253–6
  firms/companies 273
  role of 221–2
  social 578, 579
real analysis 127, 133
'real business cycle' 542, 598
real economy 83, 102, 123, 127, 129, 542
real effective exchange rate (REER) 488
real wage resistance effect 476–7
recessions 32, 53, 54, 265, 428, 451, 542, 548
  see also Great Recession
reflux 122, 159–61, 292
  principle 137–9, 152
Regional Comprehensive Economic Partnership (RCEP) 459
'The relation of home investment to unemployment' 328
relative income hypothesis 234, 255–60, 263–5
relative prices 77, 83, 124, 203, 236, 271, 333, 336–7, 475–8, 480, 487–8
rentiers 40–41, 192, 218, 220, 224, 419–21, 423, 425–9, 518
representative agents 255, 601
representative agents with rational expectations (RARE) 255, 600–601, 609
research and development (R&D)
  expenditure 472, 487
  financialization and 531–2
  underinvestment 527
'retain and reinvest strategy' 526
'Rethinking Economics' movement 30
The Return of the Master 28, 62, 615
Ricardo, D. 40, 76–7, 81, 86, 96, 106, 107, 115, 342, 407, 440, 463, 596
  classical political economics 97–9
  portrait of 118–19
  reactions against 99–100
  see also corn model (Ricardo); *On the Principles of Political Economy and Taxation* (Ricardo)
risk
  asset-backed securities 525
  banks' appreciation of 274
  credit 526
  environmental 578
  investment and 293–4, 527, 552
  Minsky on 293–4, 551, 552
  modelling 188
  principle of increasing 293, 373
  and radical uncertainty 277
  securitization and 529
  shareholder value ideology 526–7
Robinson, J. 40, 76, 98, 134, 187, 275, 399, 453–4, 579
  banana diagram 291–3
  logical time 106, 112, 579
  portrait of 466–7
Rochon, L.-P. 59, 131, 134, 172, 215, 220
Rossi, S. 215, 334–5, 498–9, 501–5, 507, 509
Russia 179, 416, 442, 448–9, 553, 563
Russian doll analogy 611
The Russian Factory in Past and Present 563

'sacrifice ratio' 221–2, 419
saltwater economics 55
Samuelson, P.A. 55, 89, 126, 275, 285, 445–6, 463
Saudi Arabia 443, 448–9
savings
  banks as intermediaries between savers and investors 156
  consumption and 236, 243, 248–9, 251, 313, 316–17
  creation and destruction of money with 138–9
  danger of excessive 137–9, 141
  definition 186
  economics approach 84
  endogenous credit-money 549
  euro area 503, 505
  finance in mainstream theory 182
  financial and banking instability 183
  impact on economic growth 234
  intentions of individuals 369–70
  investment and 189, 275, 290, 291–3, 313–14, 315, 318, 327, 357–60, 366, 367–8, 412–13, 519, 556, 608, 610
  in Kaldorian growth theory 390
  in money supply measurements 204
  natural rate of interest and 223–4, 358
  paradox of thrift 46, 84, 607
  permanent income hypothesis 234
  policy mix changes and 555–6

post-Keynesian perspective 133
supply of loanable funds 274
'virtue of' 86
see also financial savings; hoarded savings; liquid savings; negative savings
Sawyer, M. 139, 519, 521
Say, J.-B. 81, 86, 133
Say's Law 81, 86, 101, 132, 133, 134, 311, 383, 384–5
Schmitt, B. 129, 343, 348, 349, 352
Schumpeter, J.A. 46, 83, 89, 111, 113, 115, 127, 129, 197, 597, 610–11
science, economics as 34–7, 39–40, 60, 64–5
scientific realism 609–10
Seccareccia, M. 133, 141, 157, 172, 220
secondary markets 168, 178–9, 195
secular stagnation 29–30, 136, 300, 386
securitization 523, 525, 529–30
self-adjusting economy, mainstream view 541–2
The Selfish Gene 602, 603
Sen, A. 568, 570, 576, 584, 589–90
services *see* goods and services
Setterfield, M. 59, 220, 560
shadow banks 195
shareholder value 519, 526–7, 531, 532
shares
in balance-of-payments crisis 474
definition 177
dividends on 192, 194
financial intermediation 187
firms' motivation to invest 269
hedging 178, 186
industrial capitalism 189–90
private sector 138
secondary markets and 178–9, 195
speculation 194
stock options 524
trade 475
short-run aggregate supply (SRAS) 314
short-run consumption function 245–6, 251–2, 255, 257–8, 259
short-run demand function 324
short-run price theory 568
short-termism 298, 518, 527, 531–2
Simon, H. 253, 254
simple income multiplier 321–2, 323
Singapore 449
single currency 490, 495–6, 498, 502, 508
Skidelsky, R. 615–16
see also The Return of the Master
Smith, A. 40, 48, 84–5, 86, 97, 152, 256, 342, 393, 420, 422, 430, 453, 541, 543, 596
social and circular process of production 83, 92, 93, 101, 102
social capital 582
The Social Foundations of Cooperatives 563
social imbalance 585
social sciences 26, 40–43, 605, 611
social surplus 81, 83, 92, 96, 108
social well-being 423–7, 429, 576
'socialization of investment' 279–80
society 40–41, 69, 92–3, 429, 537, 580–82
Solow, R.W. 55, 57, 60–61, 63–4, 470–71, 560, 567, 569, 570, 571, 598
Solow–Swan model 568–9
South Africa 416
South Korea 381, 439, 446, 449, 453, 458, 461, 478
sovereign debt crisis 352, 495, 501, 505
Spain 367, 450–51, 502–3
special-purpose vehicle (SPV) 525, 529
speculative finance 551
spending
consumption 159, 192, 219, 242–3, 260, 274, 280–81, 292, 304, 392
credit-funded 549
cuts in 332, 545, 612
on domestically produced goods and services 390
infrastructure 580
investment 274, 280–81, 287, 289, 292, 301, 386, 389–90, 393, 548
private sector 29, 557
social 326
total 543, 546, 608
see also government expenditure; public expenditure
spending power of the population 81, 91, 111
Sraffa, P. 76–7, 97, 100–101, 102, 104, 107, 275, 328
Stability and Growth Pact 366, 502
stabilization
macroeconomic 224
output 223, 224
requirements 553
stabilization policy 59–60, 253
stagflation period (1970s) 90, 207, 386, 399, 481, 547, 598
standard of living 332
broadly based 566, 574, 575
capital accumulation and 350
CPI and 340

determinants 470
economic growth and 573, 574
measuring 568, 582–5
for sustainable development 567, 582–5
'The state of macro' 58
state role in monetary circuit 141–3
Steindl, J. 191, 192–3
Stiglitz, J.E. 34, 64, 66, 69–70, 183–4, 537
stock exchange 177, 189, 190, 277–9, 318
Stockhammer, E. 530, 605
Stolper–Samuelson theorem 445–6, 463
stove metaphor 283
structural change 270, 339, 341, 377, 381–2, 479–80, 578
structural monetary reforms 347
supermultiplier 81, 97, 104, 105, 106, 109–13
supply and demand
aggregate 387, 423
for bonds 193
classical/Keynes model 419
'cobweb theorem' 399
equilibrium between 87–8
functional income distribution 412
inflation and 333
interaction of 396, 541
law of 83, 100, 103, 187–8
in loanable funds market 274
principle of 89–90, 104, 118
shocks 91
wages and 342
wages rates and productivity growth 547
supply chains 52, 457, 462
supply-led growth 377, 382–4, 396–7
surplus principle 81, 98–9, 103, 118
sustainable development
capital stocks for 566, 568, 574, 579–82, 585, 590
definition 566
economic growth and 567–77, 579–85
estimating non-market values for 583
heterodox economics 575–9
investment for 579–82
mainstream perspective 567, 569–71, 572–3, 583
measuring new standard of living for 582–5
neoclassical model 568–9, 589
post-Keynesian sustainability model 577
'strong' and 'weak' sustainability debate 569–73
sustainable indicators 566
Swan, T. 560, 568–9
Switzerland 111, 449, 504, 522
symmetric adjustments 557–9

tableau economique fondamental see fundamental economic table
Taiwan 448–9, 461
TARGET2 system 504–8
tariffs 438, 447, 452, 453–4, 457–9, 466, 491
tax multipliers 234, 243
tax obligations 362
tax rates 357, 361, 362, 364–5
tax receipts 369
tax revenues 355, 356–7, 363, 365, 367, 370–71, 556–7, 558
taxation
changes in 235
indirect 338
progressive 244
taxes
autonomous 244
carbon 571, 573
consumption and 260, 320
disposable income 234, 316–17
distribution of stock of wealth 404
financial transactions 534
fiscal policy 326, 356, 557
government deposits and 168
income multiplier and 322–3
in monetary circuit with state 142
public expenditure mainly financed by 320
role of automatic stabilizers 363–4
see also tariffs
Taylor rule 59, 201, 208, 223, 224
technical change 393, 471
technical coefficient of production 284–7, 299
technical models 60
technical progress 110, 270, 304, 383, 393–4, 396, 398–9, 470–72
technological capability 311
technological change 91, 341, 428, 439, 456, 533, 569
technological development 569, 577
technological improvements 338, 341, 393, 444, 451–2, 524
technological innovators 461
technology
economic development pushed by 575
sharing 452–3

Solow's model and 470–71
sustainability and 569, 570, 571, 582
see also new technologies
terms of trade
changes in 442–4
definition 438
global, for crude oil and non-fuel primary commodities 443–4
textbooks
approach to finance 185–8, 191
approach to macroeconomics 595–6
place of money in 124, 152–3, 161, 169–70
Thailand 449, 478
Thatcher, M. 27, 57, 399, 599
Theory of Interest 235
theory of monetary circuit 129–41, 143
The Theory of Moral Sentiments 48
theory of output and employment 109–11, 114, 234, 594, 597
A Theory of Profits 294
theory of value and distribution 76, 84, 99, 114
Thirlwall, A.P. 304, 399, 471, 473, 479, 480, 487–8, 490–91
Thirlwall's Law 469, 470, 477–9, 484, 487, 488, 491
TINA (there is no alternative) 27, 57, 69, 123
Tobin, J. 55, 278, 569
Tobin tax 278
Tobin's Q 278–9, 299
total demand 33–4, 191, 209–10, 234, 351, 544, 546–7, 548, 553
Toward a New Economics: Essays in Post-Keynesian and Institutionalist Theory 227
trade
balance-of-payments 554
constrained growth 475, 477, 479
distributional consequences of 445–7
euro area 502, 504–5, 509
heterodox perspective 128, 143
inequality 446
investment and 456–7, 462
neoclassical perspective 124–7
new economic nationalism and 456–60
specialization by 393–4
terms of 438, 442–4
unemployment and 450–51
see also foreign trade; international trade
trade agreements 453–60
trade balance 438, 449–51
trade imbalances
in euro area 504–5, 509
between exports and imports 321
global 448–50
heterodox perspective 447–51
new economic nationalism and 456–7, 460
unemployment and 450–51
trade liberalization 438, 439, 454, 490
trade surpluses 292, 321, 448–50, 462
trade unions 32, 83, 92, 313, 325, 342, 545
trade war(s)
definition 438
Trump's 457–9
traditional analysis of inflation 333–40
criticisms 335–8
traditional mainstream economics *see* mainstream economics
traditional theory of comparative advantage 440–47, 448
static nature of 451, 466
Traite d'economie politique 86
Trans-Pacific Partnership (TPP) 457–8
A Treatise on Money 328
trend, cyclical movements around 111–13
trend output
determination of 109–10
supermultiplier and 110–11
triangular nature of money 215–17
trickle-down effect 567, 574–5, 581, 582
Triffin dilemma 512
Triffin, R. 507, 512
'truck and barter' 153
true economic development 575–9
Trump, D.J. 456, 457–9
Tugan-Baranovsky, M.I. 563
Turkey 448–9
Turrell, A.E. 66, 68–9

uncertainty 131
about future 69, 103, 110
banks 134–6, 212
engendered by Trump 458
government's stabilizing influence 64
importance for spending decisions 141–3
Keynes's conception of 33, 239–40, 277
micro- and macro- 134–6
randomness of 254
wrong expectations and 132

see also fundamental uncertainty; radical uncertainty
underconsumption 191, 543, 546, 560
underinvestment 527
unemployment
during 1970s 90, 325
in aggregate supply model 314–15
automatic stabilizers and 363, 365
case for wage cuts to reduce 595, 612
causes of 92, 111, 145, 190, 193, 308, 339, 342–3, 350–51, 537, 557
as consequence of labour market rigidities and changes in consumer demand 342–3
definition 340
in developing countries 472
economic stagnation and 191
economics approach 84
effect on wages 101
elaborated multiplier and 322–3
fiscal policy 234, 261, 265, 326, 373
frictional 332, 340–42
heterodox approach 447
income distributive effects of monetary policy and 218–19
interest rates and 360
long-period trend 112
monetarist approach to monetary union 496
as only temporary (assumption) 38
paradox of thrift 46
pessimistic banks and 141
Phillips curve 209, 213, 598
'sacrifice ratio' 499
secular stagnation and 386
self-adjusting market mechanism and 325
trade and 450–51
as usual economic condition 368
voluntary 332, 340–41
see also involuntary unemployment
unions *see* trade unions
unit labour costs 448–9, 463, 501, 547
United Arab Emirates 449
United Kingdom
balance-of-payments constrained growth 479
Brexit referendum 456, 457
budget deficits 367
Companies Acts 189
economic growth in long run 378
fallacy of composition example 606
financial and banking instability 183
financial markets 179
financialization and 521, 522
global trade imbalances 448–9
housing market 180
large financial sector 177
pound's purchasing power 336
Queen Elizabeth II 38, 62
stocks and shares 180
United States
budget deficits 367
Companies Acts 189
comparative advantage 440–42, 451
dollar's purchasing power 336
economic growth in long run 378
'exorbitant privilege' of 559
exports 320
financial and banking instability 183
financial markets 179
financialization and 519, 521–2, 524, 526, 531, 532–3
fiscal equalization mechanism 504
global trade imbalances 448–9
high unemployment rates 325
housing market 180
income disparity 393, 416–18
international monetary system 559–60
large financial sector 177
new economic nationalism 457–9
paradox of thrift 392
productivity growth rates 471
quality of life 574
sources of demand in long run 386
stocks and shares 180, 519
see also financial crises: American/subprime
United States–Mexico–Canada Agreement (USMCA) 454–5, 458–9
unsteady growth 377, 380, 382
utility maximization
behaviour 60, 312
with forward-looking agents 246
of households 82
of individuals 589, 590, 594, 596, 599
intertemporal choice and 234, 235, 236, 237
Keynes's reservations about model 239, 240
life-cycle income hypothesis 234, 247–8
marginal analysis 86
permanent income hypothesis 251
utility theory of value 125

value-added 424, 438, 460, 461–2
values
  labour 83, 97–8, 100, 106–7
  non-market 106–7
  trend 109, 111–12
Veblen, T. 190–91, 254, 256, 399, 575
'velocity of circulation' of money 204–5, 206
Verdoorn, P.J. 270, 394
Verdoorn's Law 270, 393–4, 399, 472
vicious and virtuous circles 377, 395, 399
Vines, D. 53, 56, 66–7
'virtue of saving' 86
voluntary unemployment 332, 340–41

wage inequality 446–7
wage rate(s) 91, 99, 101, 132
  average 90
  money 85, 101, 104, 107–8
  natural 99
  and productivity growth 547
  real 98
wages
  aggregate supply 323–4
  in agriculture 95, 96–7, 98–9
  banks and 134
  call for reforms 534
  cuts in 612
  determination of 412
  dual and contradictory nature of 544, 546
  economic growth and 392–3
  financialization and 531–2
  households saving 360
  income distribution and 405, 409, 412, 419, 420–21, 424–6, 428, 431, 530
  inflation and 221, 335
  in investment theories 271, 280, 283, 287, 297–8
  in Kalecki's profit equation 292
  Keynes's notion of relative real 255
  in monetary circuit with state 142
  money 101, 104, 108, 255, 263, 309, 312, 342
  nominal 209, 342–3, 449–50, 476, 547
  profit and 308–10, 335, 392–3, 608
  real 108, 221, 315, 373, 476, 547, 595
  socially necessary (natural) 81, 83, 85
  as stagnating in real terms 41
  trade imbalances and 448–50
  unemployment and 219, 342–3
Walras, L. 87, 239, 407, 541, 596
Walrasian auctioneer 91, 612
Walrasian economics 83, 87–9, 90–91, 102, 103, 113–14
Walrasian–Marshallian framework 89
wealth
  accumulation 84, 430
  definition of 402–29
  polarization of 41
  stock of 404, 405, 423–7
wealth channel 218–19
wealth distribution
  commitment to Keynesian policies 427–9
  heterodox perspective 419–27
  importance of 403
  main concepts 404–5
  mainstream economic theory of 405–13
  new macroeconomics approach 423–7
  states of income distribution 413–19
Wealth of Nations 48, 118, 152
welfare economics 399–400, 576, 584, 589
welfare state 308, 319, 322–3, 386, 413, 427
well-being
  broadly based standard of living 566
  of citizens 222–3
  economic 347, 377, 429, 584, 589
  economic growth and 567, 568, 574, 581
  poverty and 571
  sustainability and 590
  use of personal income to measure 583
  of whole population 225
  workers' 532
'wellspring of creativity' 30, 72
Werner Plans 496–8
Wills, S. 53, 56, 66–7
Wood, A. 294
workers
  capitalists and 544, 609
  distributional consequences of trade 445–7
  division of wealth and 41
  factor substitutability 271
  in Goodwin's cycles 297
  income distributive effects of monetary policy and 218–20
  installation of new equipment 272
  interactions among rentiers, entrepreneurs and 419–23, 425, 429
  involuntary unemployment 83
  in Kalecki's profit equation 292

in Keynes's notion of relative real wage 255
luxury consumption and 191
Marx on 430–31
in monetary circuit analysis 130–31, 134
risk borne by 526–7
social surplus and 92
unemployment and 342, 447, 463
well-being 532

zero lower bound 136, 359